

# Core-Periphery Organisation of Human Brain Dynamics

Danielle S. Bassett<sup>1,2,\*</sup>, Nicholas F. Wymbs<sup>5</sup>, M. Puck Rombach<sup>3</sup>, Mason A. Porter<sup>3,4</sup>, Peter J. Mucha<sup>6,7</sup>, Scott T. Grafton<sup>5</sup>

<sup>1</sup>*Department of Physics, University of California, Santa Barbara, CA 93106, USA*

<sup>2</sup>*Sage Center for the Study of the Mind, University of California, Santa Barbara, CA 93106, USA*

<sup>3</sup>*Oxford Centre for Industrial and Applied Mathematics, Mathematical Institute, University of Oxford, Oxford OX1 3LB, UK*

<sup>4</sup>*CABDyN Complexity Centre, University of Oxford, Oxford, OX1 1HP, UK*

<sup>5</sup>*Department of Psychological and Brain Sciences and UCSB Brain Imaging Center, University of California, Santa Barbara, CA 93106, USA*

<sup>6</sup>*Carolina Center for Interdisciplinary Applied Mathematics, Department of Mathematics, University of North Carolina, Chapel Hill, NC 27599, USA*

<sup>7</sup>*Institute for Advanced Materials, Nanoscience & Technology, University of North Carolina, Chapel Hill, NC 27599, USA*

As a person learns a new skill, distinct synapses, brain regions, and networks are engaged and change over time<sup>1</sup>. To better understand the dynamic processes that integrate information across a set of regions to enable the emergence of novel behaviour, we measure brain activity during motor sequencing and characterise network properties based on coherent activity between brain regions. Using recently developed algorithms<sup>2-4</sup> to detect time-evolving communities of brain regions that control learning, we find that the complex reconfiguration patterns of local communities can be described parsimoniously by the combined presence of a relatively stiff *core* of primary sensorimotor and visual regions whose connectivity changes little in time and a flexible *periphery* of multimodal association regions whose connectivity changes frequently. The separation between core and periphery changes with the duration of task practice and, importantly, is a good predictor of individual differences in learning success. This temporally defined core-periphery organisation corresponds with notions of core-periphery organisation established previously in social networks<sup>5</sup>: The geometric core<sup>6</sup> of strongly connected regions tends to also be the temporal core of stiff regions. We then show, using hypergraphs, that the core is dominated by the co-evolution of network edges, demonstrating that core regions turn on and off with the same set of partner regions over the slow time scale of learning. Our results demonstrate that an organisational structure—namely, core-periphery structure—provides a fundamental new approach for understanding how separable functional modules<sup>7-9</sup> are linked. This, in turn, enables the prediction of fundamental capacities, including the production of complex goal-directed behaviour, in humans.

Cohesive structures have long been thought to play an important role in information processing in the human brain<sup>10</sup>. At the small scale of individual neurons, temporally coherent activity supports information transfer between cells<sup>11</sup>. At a much larger scale, simultaneously active cortical areas form functional systems that enable behaviour<sup>10</sup>.

However, the question of precisely what type of cohesive organisation is present between the constituents of brain systems has been steeped in controversy<sup>12,13</sup>. It has been argued that the brain consists of independent non-interacting computational modules that are not subject to external modulation<sup>7</sup>, but this drastic view of module encapsulation has largely been dismantled by recent evidence for dynamic module interaction<sup>14</sup>. Moreover, although interactions between pairs of regions are easy to measure, the simultaneous characterisation of dynamic interactions across the whole human brain remained challenging until recent applications of network theory to neuroimaging data<sup>15</sup> (see Figure 1A). These efforts have led to enormous insights, including the establishments of links between functional brain network configuration and intelligence<sup>16</sup> and also between altered brain network organisation and disease<sup>17</sup>.

In this paper, we ask the following question: What cohesive structures in brain networks can be used to capture relevant brain dynamics that are particularly relevant for characterising skill learning that is operating over relatively long time scales (minutes to hours) compared to the dynamics of individual neurons? To answer this question, we combine new computational tools<sup>2,4</sup> for temporal networks<sup>18</sup> with ideas from social science and mathematics. Our results suggest that core-periphery organisation (see Figure 1B) is a critical property that is as important as modularity

for understanding and predicting cognition and behaviour.

### **Existence of a Temporal Core and Periphery**

We represent putative communication patterns between brain regions as a function of time (see Figure 2). This representation as a time-dependent network<sup>18</sup> allows us to examine the temporal evolution of functional connectivity (‘edges’) between anatomically distinct brain areas (‘nodes’).

To identify cohesive structures in temporally evolving brain networks, we use a form of data clustering known as ‘community detection’<sup>19</sup> to partition brain regions into groups that display similar blood oxygen level-dependent (BOLD) activity and therefore might serve distinct functions. The technique that we employ<sup>2</sup> allows the group membership of brain regions in a given subject to change over approximately 3 min time intervals (during which subjects completed an average of 10 motor sequence trials). We use the modularity index  $Q$  (defined in the SI) to estimate the strength of community structure in a network. We find that  $Q$  decreases over the 6 weeks of learning (see the SI), indicating growing interactions between communities and suggesting increasing functional integration. Additionally, the network structure of later learning contains more communities, suggesting that more putative functional modules are necessary to support specific learning functions, while also exhibiting an increased temporal variability of community membership, suggesting that brain regions have an increased freedom to communicate with multiple different communities through time.

To map the learning-dependent changes of whole-brain network structure with a more bi-



ologically insightful model, we seek to identify the roles of individual brain regions in cohesive temporally-evolving functional structures. We define the *flexibility*  $f_i$  of node  $i$  to be the number of times that it changed modular assignment over time in the network. We define a *temporal core*, *temporal bulk*, and *temporal periphery* to consist, respectively, of brain regions whose mean flexibility over participants is significantly less than, not significantly different from, or significantly more than expected in a comparable null model<sup>4</sup> (see Figure 3A and the SI for further details). We find that 68 brain regions in the frontal and temporal cortex comprise the bulk, 19 regions in the primary sensorimotor and visual cortices comprise the relatively stiff (i.e., not flexible) core, and 25 regions in multimodal association areas in visual and parietal cortex comprise the relatively flexible periphery (see Figure 3B). As discussed in the SI, the delineation of the brain into these three groups is robust to changing skill levels over the course of training and to iterative measurement (i.e., 4 MRI scanning sessions over approximately 6 weeks). The temporal core, bulk, and periphery furthermore tend to form their own communities, although this relationship appears to be modulated by learning (see SI).

### **Temporal Core-Periphery Organisation Predicts Learning**

The presence of a temporally stiff core consisting primarily of unimodal regions and a flexible periphery of multimodal cortices is consistent with the different roles of these cortices: the former are thought to perform visuomotor functions, whereas the latter perform a broader range of cognitive functions<sup>20</sup>. The ability to retrieve and rapidly execute complex motor sequences requires extensive practice. These well-learned sequences are known to be generated by our “core” areas<sup>21–24</sup>.

However, when first learning a sequence, people can use a variety of cognitive strategies supported by other brain systems (some of which are located in our periphery to augment performance)<sup>25,26</sup>. In some cases, these strategies are detrimental to skill retention<sup>27</sup>. Consequently, we hypothesised that individuals whose core-periphery organisation is well-delineated—indicating a strong separability of visuomotor and cognitive regions—would learn better than those whose core and periphery were less distinguishable from one another. Indeed, the skewness (and therefore the asymmetry of the temporal core and periphery) of flexibility over brain regions estimated on day 1 of the experiment predicted how well individuals learned in the subsequent 10 home training sessions (the Spearman rank correlation is  $\rho \approx -0.49$ , and the p-value is  $p \approx 0.027$ ) (see Figure 4A). These results suggest that successful brain function might depend on a delineation between a set of core regions whose allegiance to functional modules changes little over time and a set of periphery regions whose allegiance to functional modules is flexible through time. A similar prediction is not possible using mean regional power or pairwise coherence (see the SI).

### **Geometric Core-Periphery Organisation Correlates Strongly with Temporal Core-Periphery Organisation**

The presence of a behaviourally important core-periphery organisation based on temporal dynamics provides an opportunity to examine relative differences of underlying network architecture between the core and periphery (as well as their interactions over training). Drawing on studies of social networks<sup>5</sup>, we test whether nodes in the core are densely connected to one another and whether nodes in the periphery are sparsely connected to one another. To do this, we use a new

method<sup>6</sup> to calculate a *geometric core score*, which yields a continuous measure of core versus periphery nodes, using data from static networks that we obtained by aggregating the time series over individual time windows (i.e., in individual layers of the multilayer functional brain networks<sup>3</sup>).

We find that the geometric core score is negatively correlated with flexibility, indicating that dynamically flexible brain regions tend to also be regions that are sparsely connected to other nodes, whereas regions whose community allegiance changes little over time tend to also be regions that are strongly connected to other regions (see Figure 4A,B). That is, the temporal core of the functional brain network is strongly correlated to its geometric core. As we discuss in the SI, the negative correlation between the geometric core score and flexibility is consistently observed across participants and is robust to changes in skill level and to repeated measurement (i.e., 4 scanning sessions over approximately 6 weeks). We conclude that nodes in the temporal core have strong functional links to other nodes in the core over a very long time scale. Importantly, our demonstration of core-periphery organisation in functional brain networks parallels previous suggestions that there might be a core of highly interconnected hubs in structural brain networks<sup>28</sup>.

### **Cohesive Evolution of Brain Dynamics**

The behaviour of changing brain networks with a core-periphery organisation could take multiple forms. Over the slow time scale of motor training, communication between brain regions could either vary in strength or remain relatively constant in time. Moreover, if it is variable, then sets of connections could all change in a similar way with one another or they could vary relatively

independently (see Figure 1C).

To determine which pattern best reflects learning-dependent changes of brain dynamics on a time scale of 2–3 minutes, we use the formalism of hypergraphs to describe the cross-linking structure in dynamic human brain networks. Each hyperedge is defined by a set of functional connections (i.e., network edges) that show similar edge-weight dynamics during the learning experiment. The size (i.e., cardinality) of a hyperedge is given by the number of co-varying functional connections. As discussed in the SI, the distribution of hyperedge sizes is heavy-tailed, and the small hyperedges are located predominantly in the sensorimotor and primary visual cortices.

We construct a brain hypergraph from the hyperedges that characterise brain network evolution. To quantify a brain region’s participation in the hypergraph, we define the hypergraph node degree to be the number of hyperedges in which a node participates. We find that regions in the temporal core participate in more hyperedges than regions in the temporal periphery (see Figure 4B). Moreover, hypergraph node degree is negatively correlated with flexibility, and this relationship is robust to changes in skill level and to iterative measurement. These results demonstrate that connections incident to the brain’s core co-vary strongly with one another over 2–3 minute time scales (irrespective of whether they terminate at other core regions, bulk regions, or periphery regions.)

## **Conclusions**

We have shown how the mesoscopic organisation of functional brain networks changes over the course of learning. A key finding is that such features of a brain network provide a much better predictor than individual brain regions or small sets of regions for individual differences in subsequent ability to learn. Our results suggest that core-periphery organisation is an important and predictive component of cognitive processes that support sequential, goal-directed behaviour. They also demonstrate that during the generation of motor sequences the brain consists of a stable, densely connected, and cohesively evolving set of core regions complemented by a flexible, sparsely connected, and disparately evolving set of peripheral regions. This functional tradeoff between a core and periphery might provide balance between the stability necessary to maintain motor function by the core and the adaptivity of the periphery necessary to enable behavioural change as a function of context or strategy.

## **Methods**

Twenty healthy, right-handed participants (11 females and 9 males; mean age 24) volunteered with informed consent in accordance with the Institutional Review Board/Human Subjects Committee, University of California, Santa Barbara. Participants completed a minimum of 30 home training sessions as well as 4 fMRI sessions over the course of approximately 6 weeks. Each participant practiced a set of 10-element, visually-presented finger-movement sequences with his/her right hand. Skill levels were tuned over the course of training: 2 sequences were practiced extensively,

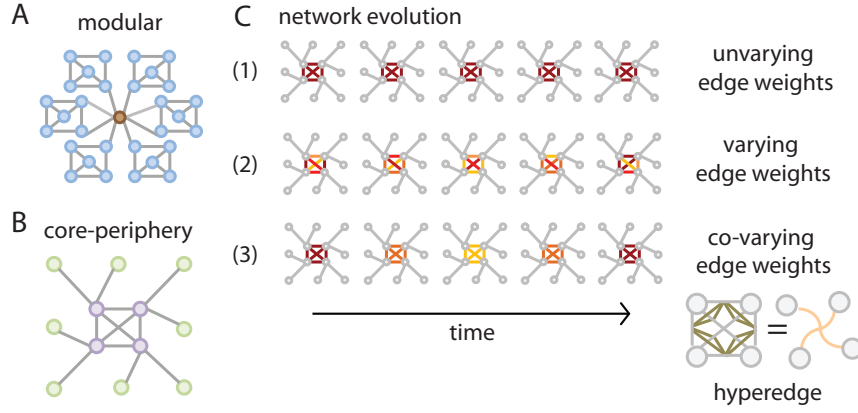


Figure 1: **Cohesive Mesoscale Structures and Dynamics** (A) A network with a modular organisation in which high-degree nodes (brown) are often found in the center of modules or bridging distinct modules that are composed mostly of low-degree nodes (blue). (B) A network with a core-periphery organisation in which nodes in the core (purple) are more densely connected with one another than nodes in the periphery are with one another (green). (C) In temporally evolving systems, edge weights (in either a modular or core-periphery network) can either (1) remain static through time, (2) vary through time, or (3) co-vary through time with one another. Edges that co-vary in weight over time are cross-linked (gold), which we represent using a hyperedge (pink) that connects all nodes that were originally connected by the co-varying edges. In the toy example shown in (3), we show a hyperedge that connects core nodes; however, hyperedges can connect nodes of any type.

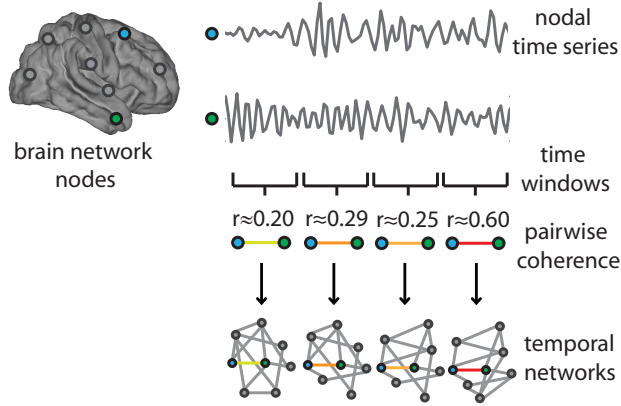


Figure 2: **Temporal Networks of the Human Brain.** We parcellate the brain into anatomical regions that can be represented as nodes in a network, and we use the coherence between functional Magnetic Resonance Imaging (fMRI) time series of each pair of nodes over a time window to determine the weight of the network edge connecting those nodes. We determine these weights separately using approximately 10 non-overlapping time windows of 2–3 min duration and thereby construct temporal networks that represent the dynamical functional connectivity in the brain.

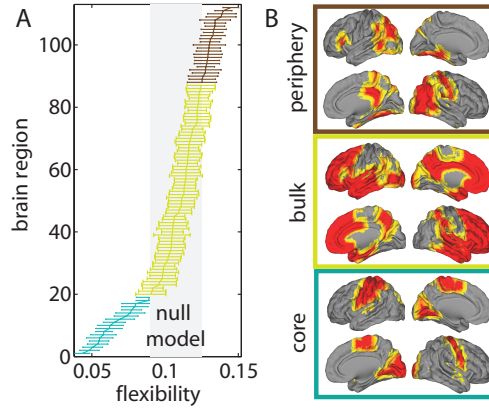


Figure 3: **Temporal Core-Periphery Organisation of the Brain** determined using fMRI signals during the performance of a simple motor learning task. (A) The core (cyan), bulk (gold), and periphery (maroon) nodes consist, respectively, of brain regions whose mean flexibility over individuals is less than, equal to, or greater than that expected in a null model (gray shaded region). (We measure flexibility based on the allegiance of nodes to putative functional modules.) Error bars indicate the standard error of the mean over individuals. (B) The anatomical distribution of regions in the core, bulk, and periphery appears to be spatially contiguous. The core contains primary sensorimotor and visual processing areas, the periphery contains multimodal association areas, and the bulk contains the remainder of the brain (and is therefore composed predominantly of frontal and temporal cortex).



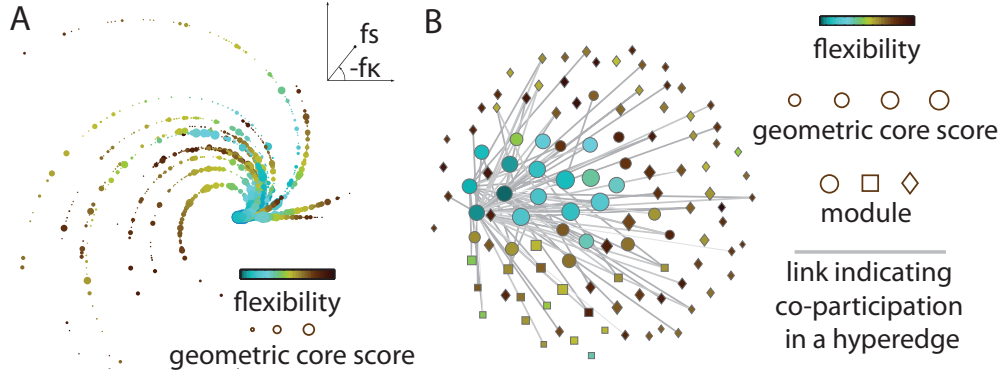


Figure 4: **Core-Periphery Organisation and Cohesive Evolution of Brain Dynamics During Learning.**

(A) The relationship between temporal and geometric core-periphery organisation is present in individual subjects, here represented as spirals in a 2-dimensional plane where data points represent brain regions located at the polar coordinates  $(fs, -f\kappa)$  where  $f$  is the flexibility of the region,  $s$  is the skewness of flexibility over all regions, and  $\kappa$  is the learning parameter (see the SI) describing each individual's relative improvement between sessions. The skewness predicts individual differences in learning (see the SI): poor learners (straighter spirals) tend to have a low skewness (short spirals) while good learners (curvier spirals) tend to have high skewness (long spirals). (B) The relationships between temporal core-periphery organisation (as measured by flexibility), geometric core-periphery organisation (as measured by geometric core score), and cohesive evolution of brain dynamics (as measured by the hypergraph node degree) can be visualised using a network representation, in which it is evident that the temporal core (i.e., low-flexibility regions of the brain) corresponds to the geometric core (i.e., regions with dense network connectivity). Hyperedges predominantly connect regions of the core with one another and with regions of the periphery but do not link peripheral regions with one another (see the SI). Importantly, this core-periphery organisation overlays community organisation. In this panel, we indicate a representative partition of the mean network over participants and experimental blocks into three communities (circles, squares, and diamonds). In this panel, we only show edges between two nodes if that pair of nodes participates in a hyperedge together in at least 60% of participants in this experiment.

2 sequences were practiced moderately, and 2 sequences were practiced minimally.

To study the predictive relationship between brain organisation and learning, we extracted brain network structure during scanning (on approximately days 1, 14, 28, and 42), and we extracted behavioural estimates of learning (using the decay constant from the single exponential fit of the duration of movement per sequence versus time) during initial home training (approximately between days 1 and 14). To study brain network organisation, we parcellated the brain into 112 cortical and subcortical regions using the Harvard-Oxford atlas. For each experimental block (2–3 min), we estimated the functional association between regions using the magnitude squared coherence of wavelet scale-two time series (0.06–0.12 Hz) to construct the temporal networks.

To identify putative functional modules, we optimised the multilayer modularity quality function<sup>2</sup> (with Newman-Girvan null models within layers) using a locally greedy Louvain<sup>29</sup> algorithm to obtain a partition of brain regions into communities for each time window. To measure changes in the composition of communities across time, we defined the *flexibility*  $f_i$  of a node to be the number of times that a node changed modular assignment throughout the set of time windows represented by the multilayer network<sup>3</sup>.

We produced brain-surface visualisations using Caret software ([brainvis.wustl.edu](http://brainvis.wustl.edu)), and we produced network visualisations using freely available MATLAB software (<http://netwiki.amath.unc.edu/VisComms/VisComms>)<sup>30</sup>.

See the SI for additional information.

**Acknowledgements** We thank Jean M. Carlson, Matthew Cieslak, John Doyle, Daniel Greenfield, and Megan T. Valentine for helpful discussions; John Bushnell for technical support; and James Fowler, Sang Hoon Lee, and Melissa Lever for comments on earlier versions of the manuscript. D.S.B. was supported by the Sage Center for the Study of the Mind, the Errett Fisher Foundation, the Templeton Foundation, the David and Lucile Packard Foundation, the Public Health Service Grant NS44393, and the Institute for Collaborative Biotechnologies through Contract W911NF-09-D-0001 from the US Army Research Office. M.A.P. and M.P.R. acknowledge a research award (#220020177) from the James S. McDonnell Foundation. M.A.P. also acknowledges a grant from the EPSRC (EP/J001759/1). P.J.M. acknowledges support from Award Number R21GM099493 from the National Institute of General Medical Sciences. The content is solely the responsibility of the authors and does not necessarily represent the official views of any of the funding agencies.

**Author Contributions** D.S.B., N.F.W., M.A.P., P.J.M., and S.T.G. formulated the project; N.F.W. and S.T.G. performed the experiments; D.S.B., N.F.W., and M.P.R. did the computations; and D.S.B., N.F.W., M.P.R., M.A.P., P.J.M., and S.T.G. wrote the manuscript.

**Competing Interests** The authors declare that they have no competing financial interests.

**Correspondence** Correspondence and requests for materials should be addressed to D.S.B. (e-mail: dbasset@physics.ucsb.edu).

1. Dosenbach, N. U. *et al.* A core system for the implementation of task sets. *Neuron* **50**, 799–812 (2006).
2. Mucha, P. J., Richardson, T., Macon, K., Porter, M. A. & Onnela, J.-P. Community structure in time-dependent, multiscale, and multiplex networks. *Science* **328**, 876–878 (2010).
3. Bassett, D. S. *et al.* Dynamic reconfiguration of human brain networks during learning. *Proc Natl Acad Sci USA* **108**, 7641–7646 (2011).
4. Bassett, D. S. *et al.* Robust detection of dynamic community structure in networks. *arXiv:1206.4358* (2012).
5. Borgatti, S. P. & Everett, M. G. Models of core / periphery structures. *Social Networks* **21**, 375–395 (1999).
6. Rombach, M. P., Porter, M. A., Fowler, J. H. & Mucha, P. J. Core-periphery structure in networks. *arXiv:1202.2684* (2012).
7. Fodor, J. A. *Modularity of Mind: An Essay on Faculty Psychology* (MIT Press, 1983).
8. Bullmore, E. & Sporns, O. Complex brain networks: Graph theoretical analysis of structural and functional systems. *Nat Rev Neurosci* **10**, 186–198 (2009).
9. Sporns, O. *Networks of the Brain* (MIT Press, 2010).
10. Saalman, Y. B., Pinsk, M. A., Wang, L., Li, X. & Kastner, S. The pulvinar regulates information transmission between cortical areas based on attention demands. *Science* **337**, 753–756 (2012).

11. Josić, K., Rubin, J., Matias, M. & Romo, R. (eds.) *Coherent Behavior in Neuronal Networks* (Springer, 2009).
12. Op de Beeck, H. P., Haushofer, J. & Kanwisher, N. G. Interpreting fMRI data: maps, modules and dimensions. *Nat Rev Neurosci* **9**, 123–135 (2008).
13. Plaut, D. C. Double dissociation without modularity: evidence from connectionist neuropsychology. *J Clin Exp Neuropsych* **17**, 291–321 (1995).
14. Fornito, A., Harrison, B. J., Zalesky, A. & Simons, J. S. Competitive and cooperative dynamics of large-scale brain functional networks supporting recollection. *Proc Natl Acad Sci USA* **109**, 12788–12793 (2012).
15. Bullmore, E. T. & Bassett, D. S. Brain graphs: graphical models of the human brain connectome. *Ann Rev Clin Psych* **7**, 113–140 (2011).
16. Langer, N. *et al.* Functional brain network efficiency predicts intelligence. *Hum Brain Mapp* **33**, 1393–406 (2012).
17. Bassett, D. S. & Bullmore, E. T. Human brain networks in health and disease. *Curr Opin Neurol* **22**, 340–347 (2009).
18. Holme, P. & Saramäki, J. Temporal Networks. *Phys Rep, in press* (2012). [arXiv:1108.1780](https://arxiv.org/abs/1108.1780).
19. Porter, M. A., Onnela, J.-P. & Mucha, P. J. Communities in networks. *Not Amer Math Soc* **56**, 1082–1097, 1164–1166 (2009).

20. Mesulam, M. M. From sensation to cognition. *Brain* **121**, 1013–1052 (1998).
21. Bischoff-Grethe, A., Goedert, K. M., Willingham, D. T. & Grafton, S. T. Neural substrates of response-based sequence learning using fMRI. *J Cogn Neurosci* **16**, 127–138 (2004).
22. Matsuzaka, Y., Picard, N. & Strick, P. L. Skill representation in the primary motor cortex after long-term practice. *J Neurophysiol* **97**, 1819–1832 (2007).
23. Picard, N. & Strick, P. L. Activation on the medial wall during remembered sequences of reaching movements in monkeys. *J Neurophysiol* **77**, 2197–2201 (1997).
24. Kennerley, S. W., Sakai, K. & Rushworth, M. F. Organization of action sequences and the role of the pre-SMA. *J Neurophysiol* **91**, 978–993 (2004).
25. Willingham, D. B., Salidis, J. & Gabrieli, J. D. Direct comparison of neural systems mediating conscious and unconscious skill learning. *J Neurophysiol* **88**, 1451–1460 (2002).
26. Destrebecqz, A. *et al.* The neural correlates of implicit and explicit sequence learning: Interacting networks revealed by the process dissociation procedure. *Learn Mem* **12**, 480–490 (2005).
27. Brown, R. & Robertson, E. Inducing motor skill improvements with a declarative task. *Nat Neurosci* **10**, 148–149 (2007).
28. Hagmann, P. *et al.* Mapping the structural core of human cerebral cortex. *PLoS Biol* **6**, e159 (2008).

29. Blondel, V. D., Guillaume, J. L., Lambiotte, R. & Lefebvre, E. Fast unfolding of community hierarchies in large networks. *J Stat Mech* P10008 (2008).
30. Traud, A. L., Frost, C., Mucha, P. J. & Porter, M. A. Visualization of communities in networks. *Chaos* **19**, 041104 (2009).

# Supplementary Material for Core-Periphery Organisation of Human Brain Dynamics

Danielle S. Bassett<sup>1,2,\*</sup>, Nicholas F. Wymbs<sup>5</sup>, M. Puck Rombach<sup>3,4</sup>,  
Mason A. Porter<sup>3,4</sup>, Peter J. Mucha<sup>6,7</sup>, Scott T. Grafton<sup>5</sup>

<sup>1</sup>Department of Physics, University of California, Santa Barbara, CA 93106, USA;

<sup>2</sup> Sage Center for the Study of the Mind,

University of California, Santa Barbara, CA 93106;

<sup>3</sup> Oxford Centre for Industrial and Applied Mathematics,

Mathematical Institute, University of Oxford, Oxford OX1 3LB, UK;

<sup>4</sup>CABDyN Complexity Centre, University of Oxford, Oxford, OX1 1HP, UK;

<sup>5</sup>Department of Psychological and Brain Sciences and UCSB Brain Imaging Center,  
University of California, Santa Barbara, CA 93106, USA;

<sup>6</sup>Carolina Center for Interdisciplinary Applied Mathematics,  
Department of Mathematics, University of North Carolina,

Chapel Hill, NC 27599, USA;

<sup>7</sup>Institute for Advanced Materials, Nanoscience & Technology,  
University of North Carolina, Chapel Hill, NC 27599, USA;

\*Corresponding author. Email address: [dbassett@physics.ucsb.edu](mailto:dbassett@physics.ucsb.edu)

October 15, 2012



# Contents

Materials and Methods . . . . .	2
Experiment and Data Acquisition . . . . .	2
Function MRI (fMRI) Imaging . . . . .	7
Network Construction . . . . .	7
Network Examination . . . . .	8
Statistics and Software . . . . .	11
Supplementary Results . . . . .	12
Flexibility Predicts Learning . . . . .	12
Dynamic Community Structure Changes with Learning . . . . .	12
Variation in Flexibility Across Brain Regions . . . . .	13
Temporal Core-Periphery Organisation . . . . .	15
Geometric Core-Periphery Organisation . . . . .	18
Relationship Between Temporal and Geometric Core-Periphery Organisation . . . . .	20
Temporal Core-Periphery Organisation and Network Co-evolution . . . . .	22
Methodological Considerations . . . . .	25
Experimental Factors . . . . .	25
Dynamic Community Detection . . . . .	27
Supplementary Discussion . . . . .	30
Network Predictions of Future Learning . . . . .	30
Dynamic Brain Networks . . . . .	30
Modularity Versus Core-Periphery Organisation . . . . .	31
Core-Periphery Organisation of Human Brain Structure and Function . . . . .	31

# Materials and Methods

## Experiment and Data Acquisition

### Experiment Setup and Procedure

Twenty-two right-handed participants (13 females and 9 males; the mean age was about 24) volunteered with informed consent in accordance with the Institutional Review Board/Human Subjects Committee, University of California, Santa Barbara. We excluded two participants from the analysis: one participant failed to complete the experiment, and the other had excessive head motion. Our analysis therefore includes twenty participants, who all had normal/corrected vision and no history of neurological disease or psychiatric disorders. They each completed a minimum of 30 behavioural training sessions as well as 3 functional Magnetic Resonance Imaging (fMRI) test sessions and a pre-training fMRI session. Training began immediately following the initial pre-training scan session. Test sessions were interspersed following every 2-week period of behavioural training, during which at least 10 training sessions were required. The training was done on personal laptop computers using a training module that was installed by the experimenter (NFW). Participants were instructed on how to run the module, which they were required to do for a minimum of 10 out of 14 days in a 2-week period. The module ran using Octave software (Version 3.2.3) and Psychtoolbox (Version 3). Participants were scanned on the first day of the experiment (scan 1), and then a second time approximately 14 days later (scan 2), once again approximately 14 days later (scan 3), and finally 14 days after that (scan 4). Not all participants were scanned exactly every two weeks; see Table 1 for details of the number of days that elapsed between scanning sessions.

We asked participants to practice a set of 10-element sequences that were presented visually using a discrete sequence production (DSP) task by generating responses to sequentially presented stimuli (see Figure 1) using a laptop keyboard with their right hand. Sequences were presented using a horizontal array of 5 square stimuli; the responses were mapped from left to right, such that the thumb corresponded to the leftmost stimulus and the smallest finger corresponded to the rightmost stimulus. A square highlighted in red served as the imperative to respond, and the next square in the sequence was highlighted immediately following each correct key press. If an incorrect key was pressed, the sequence was paused at the error and was restarted upon the generation of the appropriate key press.

Participants had an unlimited amount of time to respond and to complete each trial. All participants trained on the same set of 6 different 10-element sequences, which were presented with 3 different levels of exposure. We organised sequences so that each stimulus location was presented twice and included neither stimulus repetition (e.g., ‘11’ could not occur) nor regularities such as trills (e.g., ‘121’) or runs (e.g., ‘123’). Each training session (see Figure 2) included 2 extensively trained sequences (‘EXT’) that were each practiced for 64 trials, 2 moderately trained sequences (‘MOD’) that were each practiced for 10 trials, and 2 minimally trained sequences (‘MIN’) that were each practiced for 1 trial. (See Table 1 for details of the number of trials composed of extensively, moderately, and minimally trained sequences during home training sessions.) Each trial began with the presentation of a sequence identity cue. The purpose of the identity cue was to inform the participant what sequence they were going to have to type immediately. For example, the EXT sequences were preceded by either a cyan (sequence A) or magenta (sequence B) circle. Participants saw additional identity cues for the MOD sequences (red or green triangles) and for the MIN sequences (orange or white stars, each of which was outlined in black). No participant reported any difficulty viewing the different identity cues. Feedback was presented after every block of 10 trials; this feedback detailed the number of error-free sequences that the participant produced and the mean time it took to complete an error-free sequence.

Each fMRI test session was completed after approximately 10 home training sessions (see

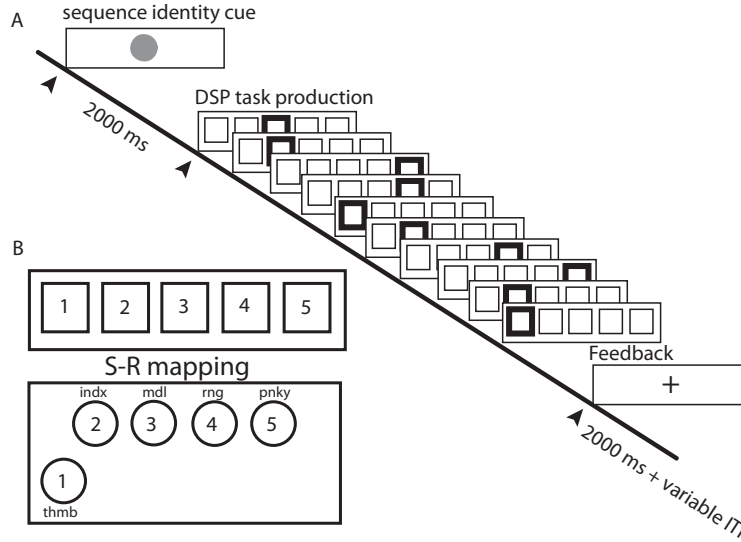


Figure 1: **Trial Structure and Stimulus-Response (S-R) Mapping.** (A) Each trial began with the presentation of a sequence identity cue that remained on screen for 2 seconds. Each of the 6 trained sequences was paired with a unique identity cue. A discrete sequence production (DSP) event structure was used to guide sequence production. The onset of the initial DSP stimulus (thick square, coloured red in the task) served as the imperative to produce the sequence. A correct key press led to the immediate presentation of the next DSP stimulus (and so on) until the 10-element sequence was correctly executed. Participants received a feedback ‘+’ to signal that a sequence was completed and to wait (approximately 0–6 seconds) for the start of the next trial. This waiting period is called the ‘inter-trial interval’ (ITI). At any point, if an incorrect key was hit, a participant would receive an error signal (not shown in the figure) and the DSP sequence would pause until the correct response was received. (B) There was a direct S-R mapping between a conventional keyboard or an MRI-compatible button box (see the lower left of the figure) and a participant’s right hand, so the leftmost DSP stimulus cued the thumb and the rightmost stimulus cued the pinky finger. Note that the button location for the thumb was positioned to the lower left to achieve maximum comfort and ease of motion.

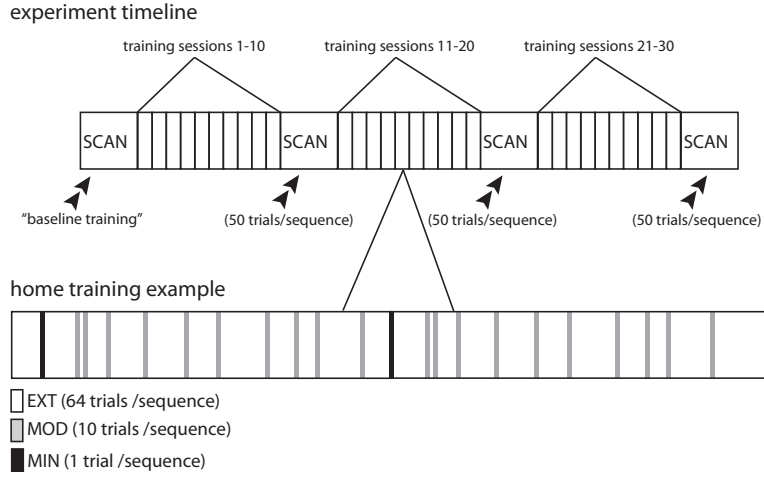


Figure 2: **Experiment Timeline.** Training sessions in the MRI scanner during the collection of blood-oxygen-level-dependent (BOLD) signals were interleaved with training sessions at home. Participants first practiced the sequences in the MRI scanner during a baseline training session (*top*). Following every approximately 10 training sessions (see Table 1), participants returned for another scanning session. During each scanning session, a participant practiced each sequence for 50 trials. Participants trained at home between the scanning sessions (*bottom*). During each home training session, participants practiced the sequences in a random order. (We determined a random order using the Mersenne Twister algorithm of Nishimura and Matsumoto [1] as implemented in the random number generator `rand.m` of MATLAB version 7.1. Each EXT sequence was practiced for 64 trials, each MOD sequence was practiced for 10 trials, and each MIN sequence was practiced for 1 trial.

Table 1 for details of the number of home practice sessions between scanning sessions), and each participant participated in 3 test sessions. In addition, each participant had a pre-training scan session that was identical to the other test scan sessions immediately prior to the start of training (see Figure 2). To familiarise participants with the task, we gave a brief introduction prior to the onset of the pre-training session. We showed the participants the mapping between the fingers and the DSP stimuli, and we explained the significance of the sequence identity cues.

To help ease the transition between each participant’s training environment to that of the scanner, padding was placed under his/her knees to maximise comfort. Participants made responses using a fiber-optic response box that was designed with a similar configuration of buttons as those found on the typical laptop used during training. See the lower left of Figure 1 for a drawing of the button box used in the experiments. The button configuration was similar to that found on a representative laptop keyboard used during home training. For instance, the center-to-center spacing between the buttons on the top row was 20 mm (compared to 20 mm from ‘G’ to ‘H’ on a MacBook Pro), and the spacing between the top row and lower left ‘thumb’ button was 32 mm (compared to 37 mm from ‘G’ to the spacebar on a MacBook Pro). The response box was supported using a board whose position could be adjusted to accommodate a participant’s reach and hand size. Additional padding was placed under the right forearm to minimise muscle strain when a participant performed the task. Head motion was minimised by inserting padded wedges between the participant and the head coil of the MRI scanner. The number of sequence trials performed during each scanning session was the same for all participants, exception for two abbreviated sessions that resulted from technical problems. In each case that scanning was cut short, participants completed 4 out of the 5 scan runs for a given session. We included data from these abbreviated sessions in this study.

Participants were tested inside the scanner with the same DSP task and the same 6 sequences that they performed during training. Participants were given an unlimited time to complete trials, though they were instructed to respond quickly but also to maintain accuracy. Trial completion was signified by the visual presentation of a fixation mark ‘+’, which remained on the screen until the onset of the next sequence identity cue. To acquire a sufficient number of events for each exposure type, all sequences were presented with the same frequency. Identical to training, trials were organised into blocks of 10 followed by performance feedback. Each block contained trials belonging to a single exposure type and included 5 trials for each sequence. Trials were separated by an inter-trial interval (ITI) that lasted between 0 and 6 seconds (not including any time remaining from the previous trial). Scan epochs contained 60 trials (i.e., 6 blocks) and consisted of 20 trials for each exposure type. Each test session contained 5 scan epochs, yielding a total of 300 trials and a variable number of brain scans depending on how quickly the task was performed. See Table 2 for details of the number of scans in each experimental block.

## Behavioural Apparatus

Stimulus presentation was controlled during training using a participant’s laptop computer, which was running Octave 3.2.4 (an open-source program that is very similar to MATLAB) in conjunction with PsychtoolBox Version 3. We controlled test sessions using a laptop computer running MATLAB version 7.1 (Mathworks, Natick, MA). We collected key-press responses and response times using a custom fiber-optic button box and transducer connected via a serial port (button box: HHSC-1  $\times$  4-L; transducer: fORP932; Current Designs, Philadelphia, PA).

	Mean	Minimum	Maximum	Standard Error
Days				
Between Scans 1 and 2	12.00	9	14	0.34
Between Scans 2 and 3	12.45	10	14	0.29
Between Scans 3 and 4	12.10	9	22	0.63
Practice Sessions				
Between Scans 1 and 2	9.70	8	10	0.14
Between Scans 2 and 3	9.75	4	14	0.44
Between Scans 3 and 4	10.05	7	13	0.32
Extensively Trained Trials				
Between Scans 1 and 2	620.80	512	640	9.40
Between Scans 2 and 3	624.00	256	896	28.57
Between Scans 3 and 4	643.20	448	832	20.48
Moderately Trained Trials				
Between Scans 1 and 2	97.00	80	100	1.46
Between Scans 2 and 3	97.50	40	140	4.46
Between Scans 3 and 4	100.50	70	130	3.20
Minimally Trained Trials				
Between Scans 1 and 2	9.70	8	10	0.14
Between Scans 2 and 3	9.75	4	14	0.44
Between Scans 3 and 4	10.05	7	13	0.32

Table 1: **Experimental Details for Behavioural Data Acquired Between Scanning Sessions.** We give the minimum, mean, maximum, and standard error of the mean over participants for the following variables: the number of days between scanning sessions; the number of practice sessions performed at home between scanning sessions; and the number of trials composed of extensively, moderately, and minimally trained sequences during home practice between scanning sessions.

## Behavioural Estimates of Learning

Our goal was to study the relationship between brain organisation and learning. To ensure independence of these two variables, we extracted brain network structure during the 4 scanning sessions, and we extracted behavioural estimates of learning during home training (approximately between days 1 and 14; see Table 1).

For each sequence, we defined the movement time (MT) as the difference between the time of the first button press and the time of the last button press during a single sequence. For the set of sequences of a single type (i.e., sequence 1, 2, 3, 4, 5, or 6), we estimated the learning rate by fitting an exponential function (plus a constant) to the MT data [2, 3] using a robust outlier correction in MATLAB:

$$MT = D_1 e^{t/\kappa} + D_2 \quad (1)$$

where  $t$  is time,  $\kappa$  is the exponential dropoff parameter (which we hereafter call the ‘learning parameter’) used to describe the early (and fast) rate of improvement, and  $D_1$  and  $D_2$  are real and positive constants.  $D_1$  is an estimate of the starting speed of the participant prior to training.  $D_2$  is an estimate of the fastest speed attainable by that participant after extended training. A negative time constant  $\kappa$  indicates a decrease in MT, which is thought to indicate that learning is occurring [4, 5]. This decrease in MT has been used to quantify learning for several decades [6, 7]. Several functional forms have been suggested for the fit of MT [8, 9], and the exponential (plus constant) is viewed as the most statistically robust choice [9]. More generally, the fitting approach that we use has the advantage of estimating the rate of learning independent of initial performance or performance ceiling.

## Functional MRI (fMRI) Imaging

### Imaging Procedures

We acquired functional MRI (fMRI) signals using a 3.0 T Siemens Trio with a 12-channel phased-array head coil. For each scan epoch, we used a single-shot echo planar imaging sequence that is sensitive to blood-oxygen-level-dependent (BOLD) contrast to acquire 37 slices per repetition time (TR of 2000 ms, 3 mm thickness, 0.5 mm gap) with an echo time (TE) of 30 ms, a flip angle of 90 degrees, a field of view (FOV) of 192 mm, and a  $64 \times 64$  acquisition matrix. Before the collection of the first functional epoch, we acquired a high-resolution T1-weighted sagittal sequence image of the whole brain (TR of 15.0 ms, TE of 4.2 ms, flip angle of 9 degrees, 3D acquisition, FOV of 256 mm, slice thickness of 0.89 mm, and  $256 \times 256$  acquisition matrix).

### fMRI Data Preprocessing

We processed and analysed functional imaging data using Statistical Parametric Mapping (SPM8, Wellcome Trust Center for Neuroimaging and University College London, UK). We first realigned raw functional data, then coregistered it to the native T1 (normalised to the MNI-152 template with a re-sliced resolution of  $3 \times 3 \times 3$  mm), and finally smoothed it using an isotropic Gaussian kernel of 8 mm full-width at half-maximum. To control for potential fluctuations in signal intensity across the scanning sessions, we normalised global intensity across all functional volumes.

### Network Construction

#### Partitioning the Brain into Regions of Interest

Brain function is characterised by a spatial specificity: different portions of the cortex emit inherently different, task-dependent activity patterns. To measure functional connectivity between different brain regions, it is common to apply a standardised structural atlas to raw fMRI data [10, 11, 12]. To study regional specificity of the functional time series, we parcellated the brain into 112 identifiable cortical and subcortical regions using the Harvard-Oxford (HO) atlas (see Table 3) installed with the FMRIB (Oxford Centre for Functional Magnetic Resonance Imaging of the Brain) Software Library (FSL; Version 4.1.1) [13, 14]. For each individual participant, we determined the regional mean BOLD time series by separately averaging across all of the voxels found in each of the 112 regions.

Within each HO-atlas region, we constrained voxel selection to those voxels that are located within an individual participant’s gray matter. To do this, we first segmented each individual participant’s T1 into white and gray matter volumes using the DARTEL toolbox supplied with SPM8. We then restricted the gray-matter voxels to those with an intensity of 0.3 or more (the maximum intensity was 1.0). Note that units are based on an arbitrary scale. We then spatially normalised the participant T1 and corresponding gray matter volume to the MNI-152 template—using the standard SPM 12-parameter affine registration from the native images to the MNI-152 template image—and resampled to 3 mm isotropic voxels. We then restricted the voxels for each HO region by using the program `fslmaths` [13, 14] to include only those voxels found in the individual participant gray-matter template.

### Wavelet Decomposition

Brain function is also characterised by a frequency specificity. Different cognitive and physiological functions are associated with different frequency bands, which can be investigated using wavelets. Wavelet decompositions of fMRI time series have been applied extensively in both resting-state and task-based conditions [15, 16]. In both cases, they provide sensitivity for the detection of small signal changes in non-stationary time series with noisy backgrounds [17].

In particular, the maximum-overlap discrete wavelet transform (MODWT) has been used extensively in connectivity investigations of fMRI [18, 19, 20, 21, 22, 23]. Accordingly, we used MODWT to decompose each regional time series into wavelet scales corresponding to specific frequency bands [24].

We were interested in quantifying high-frequency components of an fMRI signal, correlations between which might be indicative of cooperative temporal dynamics of brain activity during a task. Because our sampling frequency was 2 seconds (1 TR = 2 sec), wavelet scale one provided information on the frequency band 0.125–0.25 Hz and wavelet scale two provided information on the frequency band 0.06–0.125 Hz. Previous work has indicated that functional associations between low-frequency components of the fMRI signal (0–0.15 Hz) can be attributed to task-related functional connectivity, whereas associations between high-frequency components (0.2–0.4 Hz) cannot [25]. This frequency specificity of task-relevant functional connectivity is likely due at least in part to the hemodynamic response function, which might act as a noninvertible band-pass filter on underlying neural activity [25]. Consistent with our previous work [26], we examine wavelet scale two, which is thought to be particularly sensitive to dynamic changes in task-related functional brain architecture.

### Construction of Dynamic Networks

For each brain region, we extracted the wavelet coefficients of the mean time series in temporal windows given by trial blocks (of approximately 60 TRs; see Table 2). The leftmost temporal boundary of each window was equal to the first TR of an experimental trial block, and the rightmost boundary was equal to the last TR in the same block. We thereby extracted block-specific data sets from the EXT, MOD, and MIN sequences (with 6–10 blocks of each sequence type; see Table 2 for details of the number of blocks of each sequence type) for each of the 20 participants participating in the experiment and for each of the 4 scanning sessions.

For each block-specific data set, we constructed an  $N \times N$  adjacency matrix  $\mathbf{W}$  representing the complete set of pairwise functional connections present in the brain during that window in a given participant and for a given scan. Note that  $N = 112$  is the number of brain regions in the full brain atlas (see the earlier section on “Partitioning the Brain into Regions of Interest” for further details). To quantify the weight  $W_{ij}$  of functional connections between regions labelled  $i$  and  $j$ , we used the magnitude squared spectral coherence as a measure of nonlinear functional association between any two wavelet coefficient time series (consistent with our previous study [26]). In using the coherence, which has been demonstrated to be useful in the context of fMRI neuroimaging data [25], we were able to measure frequency-specific linear relationships between time series.

To examine dynamic changes in functional brain network architecture during learning, we constructed multilayer networks by considering the set of  $L$  adjacency matrices constructed from consecutive blocks of a given sequence type (EXT, MOD, or MIN) in a given participant and scanning session. We combined the matrices in each set separately to form a rank-3 adjacency tensor  $\mathbf{A}$  per sequence type, participant, and scan. Such a tensor can be used to represent a time-dependent network structure [27, 26]. In the following sections, we describe a variety of diagnostics that can be used to characterise these multilayer structures.

## Network Examination

### Dynamic Community Detection

Community detection [28, 29] can be used to identify putative functional modules (i.e., sets of brain regions that display similar trajectories through time). One such technique is based on the optimisation of the modularity quality function [30, 31, 32]. This allows one to identify groups that consist of nodes that have stronger connections among themselves than they do to



nodes in other groups [28]. Recently, the modularity quality function has been generalised so that one can consider time-dependent or multiplex networks using *multilayer modularity* [27]

$$Q = \frac{1}{2\mu} \sum_{ijlr} \{ (A_{ijl} - \gamma_l P_{ijl}) \delta_{lr} + \delta_{ij} \omega_{jlr} \} \delta(g_{il}, g_{jr}), \quad (2)$$

where the adjacency matrix of layer  $l$  has components  $A_{ijl}$ , the element  $P_{ijl}$  gives the components of the corresponding matrix for a null model,  $\gamma_l$  is the structural resolution parameter of layer  $l$ , the quantity  $g_{il}$  gives the community (i.e., ‘module’) assignment of node  $i$  in layer  $l$ , the quantity  $g_{jr}$  gives the community assignment of node  $j$  in layer  $r$ , the parameter  $\omega_{jlr}$  is the connection strength—i.e., ‘interlayer coupling parameters’, which constitute a set of *temporal resolution parameters* if one is using the adjacency tensor  $\mathbf{A}$  to represent a time-dependent network—between node  $j$  in layer  $r$  and node  $j$  in layer  $l$ , the total edge weight in the network is  $\mu = \frac{1}{2} \sum_{jr} \kappa_{jr}$ , the strength of node  $j$  in layer  $l$  is  $\kappa_{jl} = k_{jl} + c_{jl}$ , the intra-layer strength of node  $j$  in layer  $l$  is  $k_{jl}$ , and the inter-layer strength of node  $j$  in layer  $l$  is  $c_{jl} = \sum_r \omega_{jlr}$ . We employ the Newman-Girvan null model by letting

$$P_{ijl} = \frac{k_{il} k_{jl}}{2m_l}, \quad (3)$$

where  $m_l = \frac{1}{2} \sum_{ij} A_{ijl}$  is the total edge weight in layer  $l$ . We let  $\omega_{jlr} \equiv \omega = \text{constant}$  for neighbouring layers (i.e., when  $|l-r| = 1$ ) and  $\omega_{jlr} = 0$  otherwise. We also let  $\gamma_l = \gamma = \text{constant}$ . In the main text, we report results for  $\omega = 1$  and  $\gamma = 1$ , and we evaluate the dependence of our results on  $\gamma$  and  $\omega$  in the ‘Results’ section of this supplementary document.

Optimisation of multilayer modularity (2) yields a partition of the brain regions into communities for each time window. To measure changes in the composition of communities across time (i.e., across experimental blocks), we defined the *flexibility*  $f_i$  of a node  $i$  to be the number of times that a node changed modular assignment throughout the set of time windows represented by the multilayer network [26]. We normalised this number by the total number of changes that were possible (i.e., by the number of consecutive pairs of layers in the multilayer framework). We then defined the flexibility of the entire network as the mean flexibility over all nodes in the network:  $F = \frac{1}{N} \sum_{i=1}^N f_i$ .

### Identification of Temporal Core, Bulk, and Periphery

We found that different brain regions have different flexibilities. To determine whether a particular brain region was more or less flexible than expected, we constructed a *nodal null model*, which can be used to probe the individual roles of nodes in a network. We rewired the ends of the multilayer network’s inter-layer edges (which connect nodes in one layer to nodes in another) uniformly at random. After applying the associated permutation, an inter-layer edge can, for example, connect node  $i$  in layer  $t$  with node  $j \neq i$  in layer  $t+1$  rather than being constrained to connect each node  $i$  in layer  $t$  with itself in layer  $t+1$ .

We considered 100 different rewirings to construct an ensemble of 100 nodal null-model multilayer networks for each single multilayer network constructed from the original brain data. We then estimated the flexibility of each node in each nodal null-model network. We created a distribution of expected mean nodal flexibility values by averaging flexibility over 100 rewirings and the 20 participants. We similarly estimated the mean nodal flexibility of the original brain data by averaging flexibility over the 20 participants and 100 optimisations. (We optimised multilayer modularity using a Louvain locally greedy method [33, 34]. This procedure is not deterministic, so different runs of the optimisation procedure can yield slightly different partitions of a network.) We considered a region to be a part of the temporal ‘core’ if its mean nodal flexibility was below the 2.5% confidence bound of the null-model distribution, and we considered a region to be a part of the temporal ‘periphery’ if its mean nodal flexibility was above

the 97.5% confidence bound of the null-model distribution. Finally, we considered a region to be a part of the temporal ‘bulk’ if its mean nodal flexibility was between the 2.5% and 97.5% confidence bounds of the null-model distribution.

### Geometric Core-Periphery Organisation

To estimate the geometric core-periphery organisation of the (static) networks defined by each experimental block (i.e., for each layer of a multilayer network), we used the method that was recently proposed in Ref. [35]. This method results in a ‘core score’ (which constitutes a centrality measure) for each node that indicates where it lies on a continuous spectrum of roles between core and periphery. This method has numerous advantages over previous formulations used to study core-periphery organisation. In particular, it can identify multiple geometric cores in a network, which makes it possible to take multiple cores into account and in turn enables one to construct a detailed description of geometric core-periphery organisation by ranking the nodes in terms of how strongly they participate in different possible cores. Importantly, the continuous nature of the measure removes the need to use an artificial dichotomy of being strictly a core node versus strictly a periphery node.

In this method, we consider a vector  $C$  with non-negative values, and we let  $C_{ij} = C_i \times C_j$ , where  $i$  and  $j$  are two nodes in an  $N$ -node network. We seek a core vector  $C$  that satisfies the normalisation condition

$$\sum_{i,j} C_i C_j = 1$$

and is a permutation of the vector  $C^*$  whose components specify the *local (geometric) core values*

$$C_m^* = \frac{1}{1 + \exp\{-(m - N\beta) \times \tan(\pi\alpha/2)\}}, \quad m \in \{1, \dots, N\}. \quad (4)$$

We seek a permutation that maximises the *core quality*

$$R = \sum_{i,j} A_{ij} C_i C_j, \quad (5)$$

This method to compute core-periphery organisation has two parameters:  $\alpha \in [0, 1]$  and  $\beta \in [0, 1]$ . The parameter  $\alpha$  sets the sharpness of the boundary between the geometric core and the geometric periphery. The value  $\alpha = 0$  yields the fuzziest boundary, and  $\alpha = 1$  gives the sharpest transition (i.e., a binary transition): as  $\alpha$  varies from 0 to 1, the maximum slope of  $C^*$  varies from 0 to  $+\infty$ . The parameter  $\beta$  sets the size of the geometric core: as  $\beta$  varies from 0 to 1, the number of nodes included in the core varies from  $N$  to 0. One now has the choice of either taking into account the local core scores of a node for a set of  $(\alpha, \beta)$  coordinates sampled from  $[0, 1] \times [0, 1]$  (where one weighs each choice by its corresponding value of  $R$ ) or one can take into account only the score for particular choices of  $(\alpha, \beta)$ .

### Network Co-variation: Hypergraphs

To examine the co-variation of network structure over time, we use the formalism of hypergraphs[36] to describe the sets of edges whose weights are correlated over time. For a given set of networks, we calculate the  $E \times E$  adjacency matrix  $\mathbf{X}$  whose element  $X_{ab}$  gives the Pearson correlation coefficient between the time series of weights for edge  $a$  and the time series of weights for edge  $b$ . Note that  $E = N(N - 1)/2$  is the total number of edges per layer of the original multilayer network.

Given the very large number of tests that this procedure entails, we threshold the edge-edge correlation matrix  $\mathbf{X}$  to retain only those connections that are significant. That is, we estimate the p-value associated with the Pearson coefficient  $r$  for each edge-edge correlation. Using a

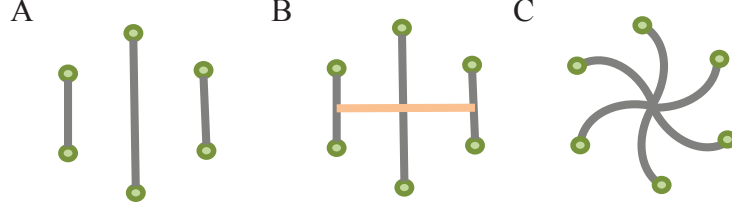


Figure 3: **Construction of Hyperedges.** A set of node-node edges [green lines; (*A*)] with strong pairwise temporal correlations among the edges [peach line; (*B*)] forms a hyperedge (*C*).

false positive correction, we binarise  $\mathbf{X}$  by setting elements to 1 when their associated p-value is less than  $2/[E(E-1)]$  and setting all other elements to 0. We use  $\xi$  to denote the binarised version of  $\mathbf{X}$ .

We examined the structure of the co-variation network by identifying the set of connected components in the adjacency matrix  $\xi$ . Each connected component represents a hyperedge, which consists of a set of edges that are connected by significant temporal correlations (see Figure 3). This yields a hypergraph consisting of the set of hyperedges identified in  $\xi$ . We define the *degree* of a node (i.e., a brain region) in the hypergraph to be equal to the number of hyperedges associated with that node.

## Statistics and Software

We performed all data analysis and statistical tests in MATLAB. We performed the dynamic community detection procedure using freely available MATLAB code [33] that optimises multilayer modularity using a Louvain locally greedy algorithm [34].

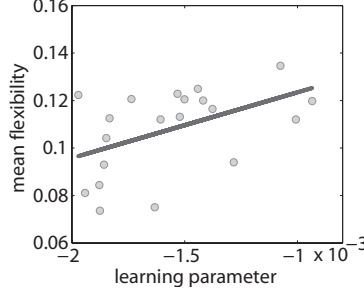


Figure 4: **Flexibility of Dynamic Community Structure Predicts Learning.** Individual flexibility, averaged over 112 brain regions and 100 multilayer modularity optimisations for extensive sequence (EXT) trial blocks, is positively correlated with the learning parameter  $\kappa$  (which we use to measure learning) of EXT sequences during the subsequent 10 home training sessions: the Pearson correlation coefficient is  $r \approx 0.48$  and the p-value is  $p \approx 0.030$ . Individuals displaying smaller mean flexibility on day 1 learned better during the subsequent 10 days of practice.

## Supplementary Results

### Flexibility Predicts Learning

In a recent study [26], we developed a dynamic network diagnostic to quantify the flexibility  $f_i$  of a brain region’s allegiance to different functional modules. In Ref. [26], we showed that the mean flexibility  $F$  over all brain regions in a 112-parcel whole brain atlas could be used to predict the amount of learning in a subsequent experimental session. Here we examine flexibility in the context of extended learning using a DSP task.

To examine the relationship between flexibility and learning, we confine ourselves to the two EXT (i.e., extensively trained) sequences, in which learning occurs more rapidly than in MOD and MIN sequences. We therefore estimate flexibility from the multilayer networks constructed from blocks of the two EXT sequences in the first scanning session. We separately estimate learning using the learning parameter  $\kappa$  in home training sessions 1–10, which take place before scanning session 2. We find that flexibility of networks derived from the two EXT sequences on day 1 are correlated with learning of those two sequences in the subsequent 10 days of home training (see Figure 4A). Importantly, the temporal separation of the data from the scanning session (which we used to estimate brain flexibility) and the home training (which we used to estimate learning) ensures that this correlation is predictive: flexibility on day 1 predicts learning in the subsequent 10 days of practice.

### Dynamic Community Structure Changes with Learning

In addition to its predictive role in learning, community structure changes over scanning sessions and is modulated by the extent of training (EXT, MOD, and MIN). In Figure 5, we show multilayer modularity, the number of communities, and mean flexibility for all 4 scanning sessions over the 3 levels of training. Modularity appears to be dominated by the effect of training: variation across EXT, MOD, and MIN is larger than variation across scanning sessions. Multilayer networks composed of EXT blocks have higher  $Q$  values than networks composed of blocks of sequences with less training. Conversely, the number of communities appears to be dominated by the effect of scanning session: variation across scans 1, 2, 3, and 4 is larger than variation across levels of training, and the results from EXT and MOD networks are not significantly distinguishable from one another. The number of communities increases with scanning session

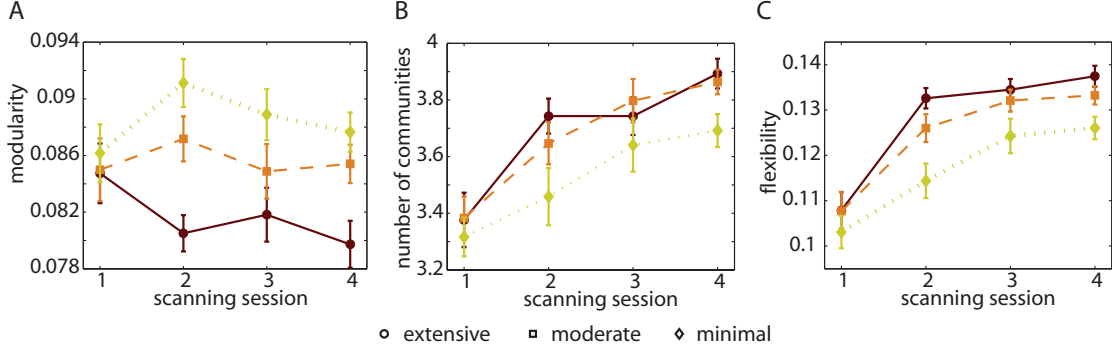


Figure 5: **Dynamic Network Diagnostics Change with Learning.** (A) Modularity, (B) number of communities, and (C) mean flexibility calculated for blocks of extensively trained (EXT; maroon circles, solid lines), moderately trained (MOD; orange squares, dashed lines), and minimally trained (MIN; gold diamonds, dotted lines) sequences over the 4 scanning sessions. We average the values for each diagnostic over the 100 multilayer modularity optimisations, and we average flexibility over the 112 brain regions (in addition to averaging over the 100 optimisations). Error bars indicate standard error of the mean over participants.

and is lowest for MIN networks. The flexibility appears to be modulated by both scanning session (increasing from scan 1 to scan 4) and level of training (decreasing from EXT to MIN).

### Variation in Flexibility Across Brain Regions

The mean flexibility over participants varied over brain regions, ranging from approximately 0.04 to approximately 0.14 (see Figure 6A). The distribution of flexibility across brain regions is decidedly non-Gaussian and seems to be bimodal: the majority of brain regions have relatively high flexibilities, but there is a left-heavy tail of regions (including a small peak) with low values of flexibility. To quantitatively characterise the distribution of flexibility over brain regions, we calculated the second (variance; see Figure 6B), third (skewness; see Figure 6C), and fourth (kurtosis; see Figure 6D) central moments of the distribution to supplement our calculation of the first moment (mean; see Figure 4). Although the variance is not significantly related to learning (the Spearman rank correlation is  $\rho \approx 0.159$ , and the p-value is  $p \approx 0.500$ ), the third and fourth moments of the distribution are significant:  $\rho \approx -0.479$  and  $p \approx 0.033$  for the skewness and  $\rho \approx -0.497$  and  $p \approx 0.027$  for the kurtosis. Importantly, these findings are predictive: the flexibility of EXT multilayer networks on scanning day 1 predicted learning over the subsequent 10 home training sessions.

To interpret these findings, we note that the skewness of the distribution is a measure of the asymmetry of the distribution, and the positive values that we observed indicate that the distributions from all participants are skewed to the right. The negative correlation between skewness and learning suggests that individuals with more skewness of flexibility over brain regions learn better than individuals with less. The kurtosis of the distribution is a combined measure of both its peakedness [37] and its bimodality [38], and it is sometimes construed as a measure of the extent that a distribution is prone to outliers. The kurtosis values that we observed vary between 2.5 and 5, which includes the value of 3 that is expected for a Gaussian distribution. The negative correlation between kurtosis and learning suggests that individuals with higher kurtosis learn better than individuals with lower kurtosis. Considered in the context of the results shown in Figure 4, we conclude that individuals with smaller mean flexibility and a more skewed but less kurtotic distribution of flexibility over brain regions seem to learn better.

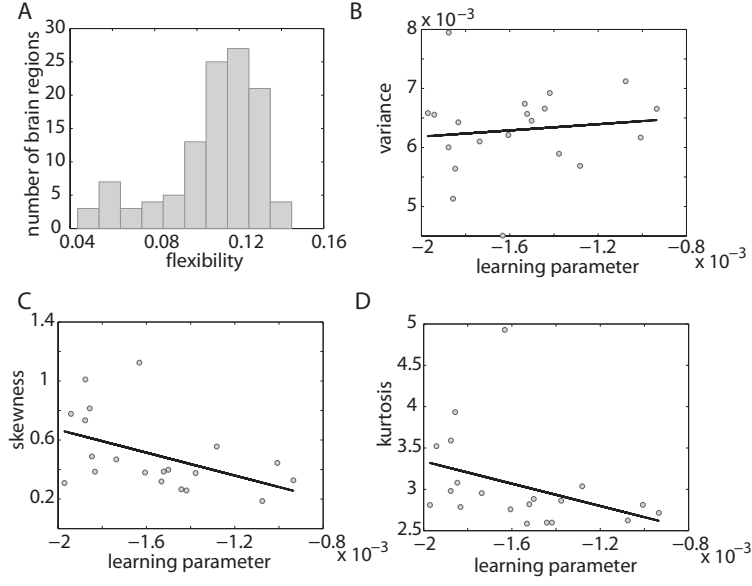


Figure 6: **Distribution of Flexibility and Learning.** (A) The distribution of flexibility across brain regions displays a left-sided tail and also appears to have two peaks: a small peak for small flexibility and a much larger peak for large flexibility. We averaged the flexibility values over the 100 multilayer modularity optimisations and the 20 participants. The (B) variance, (C) skewness, and (D) kurtosis of flexibility over brain regions averaged over the 100 multilayer modularity optimisations as a function of the learning parameter  $\kappa$  (which we use to measure learning).

## Temporal Core-Periphery Organisation

### Defining the Temporal Core and Temporal Periphery

To determine the significance of a brain region’s variation in flexibility, we compared the flexibility of brain regions in the real multilayer network to that expected in a nodal null model. We can define a temporal core, bulk, and periphery (see Figure 2A in the main manuscript). The core is the set of regions whose flexibility is significantly less than expected in the null model; the periphery is the set of regions whose flexibility is significantly greater than expected in the null model; and the bulk consists of all remaining regions.

We show the anatomical locations of the temporal core, temporal bulk, and temporal periphery in Figure 2B of the main manuscript. The relatively stiff core is composed predominantly of primary sensorimotor regions in both left and right hemispheres. The more flexible periphery is composed predominantly of multimodal regions—including inferior parietal, intraparietal sulcus, temporal parietal junction, inferotemporal, fusiform gyrus, and visual association areas. The bulk contains the remainder of the cortical and subcortical areas—including large swaths of frontal, temporal, and parietal cortices. The separation of a temporally stiff core of unimodal regions that predominantly process information from single sensory modality (e.g., vision, audition, etc.) and a flexible periphery of multimodal cortices that process information from multiple modalities is consistent with existing understanding of the association of multimodal cortex with the binding of different types of information and the performance of a broad set of functions [39].

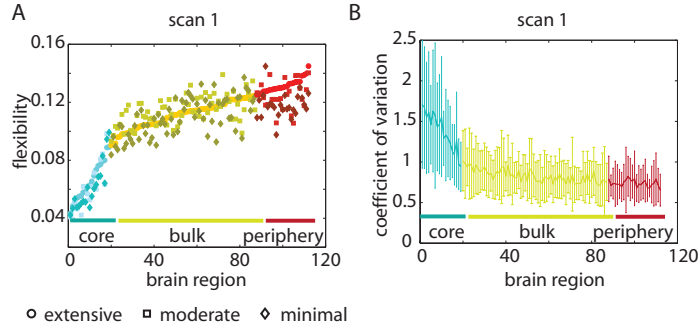
### Reliability of Temporal Core-Periphery Organisation

A brain region’s role in the temporal core, bulk, and periphery is surprisingly robust. Regions identified as part of the core, bulk, or periphery in multilayer networks constructed from the EXT blocks in scanning session 1 have similar flexibilities in the other two levels of training (MOD and MIN; see Figure 7A) for the same scanning session. To quantify the variability of a brain region’s flexibility, we calculated the coefficient of variation (CV) of flexibility over the 100 optimisations and the 3 levels of training (see Figure 7B). The CV is defined as  $CV = \sigma/\mu$ , where  $\sigma$  is the standard deviation of a given sample and  $\mu$  is its mean. We observe that the variabilities over optimisations and scans (i.e., CV) and over participants (i.e., error bars) are largest in regions designated as part of the temporal core and smallest in regions designated as part of the temporal periphery.

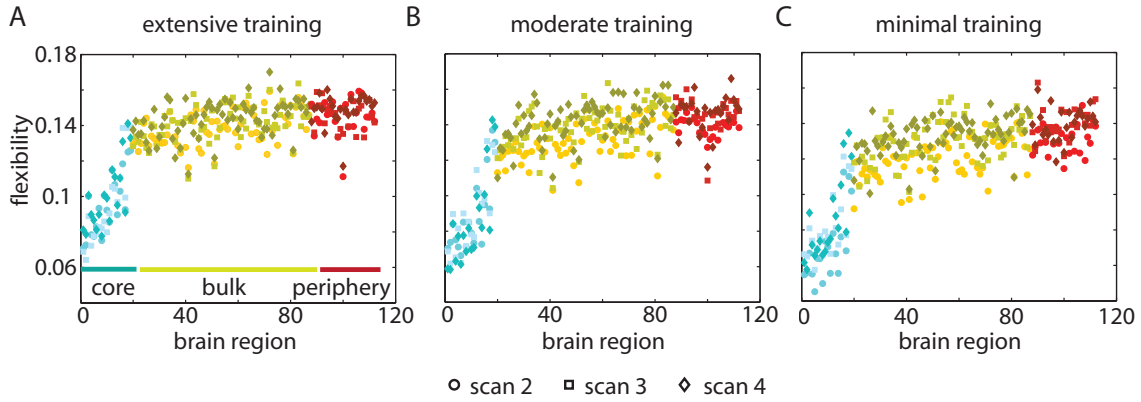
In addition, regional flexibility is also conserved over the other 3 scanning sessions over the 6-week period. Observe in Figure 8 that regions identified as part of the temporal core in multilayer networks constructed from the EXT blocks in scanning session 1 exhibit small flexibility for all other scanning sessions and for all 3 training levels (EXT, MOD, and MIN). Regions in the temporal bulk and temporal periphery exhibit a similar amount of flexibility to one another.

### Temporal Core-Periphery Organisation and Learning

In light of these data and calculations, we can reinterpret the results illustrated in Figure 6 in the following way. We had previously used the kurtosis of the distribution of flexibility over brain regions to attempt to quantify bimodality in the distribution of flexibility over brain regions. Our subsequent calculations suggest that kurtosis seems (in essence) to be measuring the separation between the temporal core and the temporal periphery. We note that the separation in mean flexibility between the relatively stiff core and flexible periphery varies over individuals and that individuals with narrower separation between temporal core and temporal periphery learn better in the subsequent 10 home training sessions than individuals with greater separation between temporal core and temporal periphery. However, we had also previously noted that



**Figure 7: Reliability of Temporal Core-Periphery Structure.** The temporal core (cyan), bulk (gold), and periphery (maroon) of dynamic networks defined by the flexibility of trial blocks in which participants practiced sequences that would eventually be extensively trained. (A) The flexibility of the temporal core, bulk, and periphery averaged over the 100 multilayer modularity optimisations and 20 participants for blocks composed of extensively trained (EXT; circles), moderately trained (MOD; squares), and minimally trained (MIN; diamonds) sequences. Colour darkness of data points indicates scanning session with darker colours indicating scan 1 and lighter colours indicating scan 4. (B) The coefficient of variation of flexibility calculated over the 100 optimisations and 3 sequence types for all brain regions. Error bars indicate the standard error of the mean CV over participants. Both panels use data from scanning session 1 on day 1 of the experiment (which is prior to home training).



**Figure 8: Temporal Core-Periphery Organisation Over 42 Days.** The temporal core (cyan), bulk (gold), and periphery (maroon) of dynamic networks defined by the trial blocks in which participants practiced sequences that would eventually be (A) extensively trained, (B) moderately trained, and (C) minimally trained for data from scanning sessions 2 (after approximately 2 weeks of training; circles), 3 (after approximately 4 weeks of training; squares), and 4 (after approximately 6 weeks of training; diamonds). Colour darkness of data points indicates scanning session with darker colours indicating scan 1 and lighter colours indicating scan 4.



individuals whose flexibility was more skewed over brain regions learned better than those whose flexibility over brain regions was less skewed. Our subsequent calculations suggest that the skewness seems (in essence) to be measuring the presence of—rather than the separation between—the temporal core and the temporal periphery. We therefore note that individuals with a strong temporal core and temporal periphery but a smooth transition between them seem to learn better than individuals with a weak temporal core and temporal periphery but a sharper transition between them.

These results suggest that successful brain function might depend on a delicate balance between the delineation into a set of core regions whose allegiance to functional modules changes little over time and a set of periphery regions whose allegiance to functional modules is flexible through time (and also on the smoothness of the transition between these two types of regions).

### Relationship Between Temporal Core-Periphery Organisation and Community Structure

The division of the whole brain network into core, bulk, and periphery has surprising similarities to its division into communities: that is, core-periphery organisation appears to map to some extent to community structure. We first noted this similarity when we examined community structure in an object we call the *mean coherence matrix*. We defined the mean coherence matrix  $\bar{\mathbf{A}}$  to contain elements  $\bar{A}_{ij}$  equal to the mean coherence between nodes  $i$  and  $j$  over participants and EXT blocks on day 1 of the experiment. We determined the community structure of this mean coherence matrix by optimising the single-layer modularity quality function [30, 40, 31, 28, 29]:

$$Q_{\text{single-layer}} = \sum_{ij} \left[ \bar{A}_{ij} - \frac{k_i k_j}{2m} \right] \delta(g_i, g_j), \quad (6)$$

where node  $i$  is assigned to community  $g_i$ , node  $j$  is assigned to community  $g_j$ , the Kronecker delta  $\delta(g_i, g_j) = 1$  if  $g_i = g_j$  and it equals 0 otherwise,  $k_i$  is the strength of node  $i$ , and  $m$  is the mean strength of all nodes in the network. After optimising this single-layer quality function 100 times, we constructed a representative partition [41] from the set of 100 partitions, each partition arising from a single optimisation (see Figure 9A). One community in this representative partition appeared to have high connectivity to the other two communities: nodes in this first community had edges with strong weights to nodes in the other two communities indicating a high coherence in BOLD time series. This behaviour is consistent with the behaviour expected from a “core”. A second community in this representative partition appeared to have low connectivity to the other two communities: nodes in this community had edges with weak weights to nodes in the other two communities, indicating a low coherence in BOLD time series. This behaviour is consistent with the behaviour expected from a “periphery”. However, because *mean matrices* constructed by averaging correlation-based matrices often do not adequately represent the geometric or topological structure of the individual matrices [42], we validated this observation by instead testing for a relationship between the temporal core-periphery organisation and community structure at the level of individual participants.

A subdivision of the brain into a temporal core, bulk, and periphery regions gives a partition of the functional brain network. We label this partition using the Greek letter  $\nu$ , and we test for similarities between this partition and algorithmic partitions, which we label using  $\eta$ , into communities for each participant, block, and optimisation using the z-score of the Rand coefficient [43]. For each pair of partitions  $\nu$  and  $\eta$ , we calculate the Rand z-score in terms of the total number of node pairs  $M$  in the network, the number of pairs  $M_\nu$  that are in the same community in partition  $\nu$  but not in the partition  $\eta$ , the number of pairs  $M_\eta$  that are in the same community in partition  $\eta$  but not in  $\nu$ , and the number of node pairs  $w_{\nu\eta}$  that are assigned to the same community both in partition  $\nu$  and in partition  $\eta$ . The z-score of the Rand

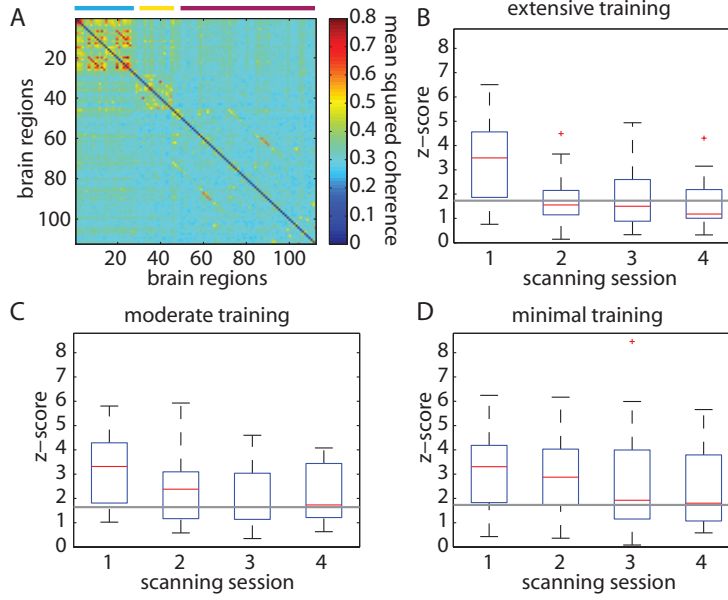


Figure 9: **Relationship Between Temporal Core-Periphery Organisation and Community Structure.** (A) Mean coherence matrix over all EXT blocks from all participants on scanning day 1. The coloured bars on the top of the matrix indicate the 3 communities identified from the representative partition. The mean partition similarity  $z_i$  over all participants for blocks of (B) extensively, (C) moderately, and (D) minimally trained sequences for all 4 scanning sessions over the approximately 6 weeks of training. The horizontal gray lines in panels (B-D) indicate the  $z_i$  value corresponding to a right-tailed p-value of 0.05.

coefficient allows us to compare partitions  $\eta$  and  $\nu$ , and it is given by the formula

$$z_{\nu\eta} = \frac{1}{\sigma_{w_{\nu\eta}}} w_{\nu\eta} - \frac{M_\nu M_\eta}{M}, \quad (7)$$

where  $\sigma_{w_{\nu\eta}}$  is the standard deviation of  $w_{\nu\eta}$ . Let the *mean partition similarity*  $z^i$  denote the mean value of  $z_{\nu\eta}$  over all partitions  $\eta$  (i.e., for all blocks and all optimisations) for participant  $i$ .

We find that communities identified by the optimisation of the multilayer modularity quality function (2) have significant overlap with the delineation into temporal core, bulk, and periphery during early learning; see Figure 9B-D. The mean values of  $z^i$  over participants indicate that there is a significant similarity between the partitions into modules and the partitions into core, bulk, and periphery for blocks of extensively, moderately, and minimally trained sequences on scanning day 1. This similarity between community structure and temporal core-periphery organisation is also evident for blocks of moderately and minimally trained sequences practiced during later scanning sessions. These results underscore the fact that core-periphery organisation can be consistent with community structure, and they also suggest that the relationship between these two types of mesoscopic organisation can be modulated by learning.

## Geometric Core-Periphery Organisation

Given our demonstration that there exists a temporal core in dynamic brain networks, it is important to ask what role such core regions play in individual network layers (i.e., in aggregations of the coupled time series over individual time windows into static networks) of the multilayer

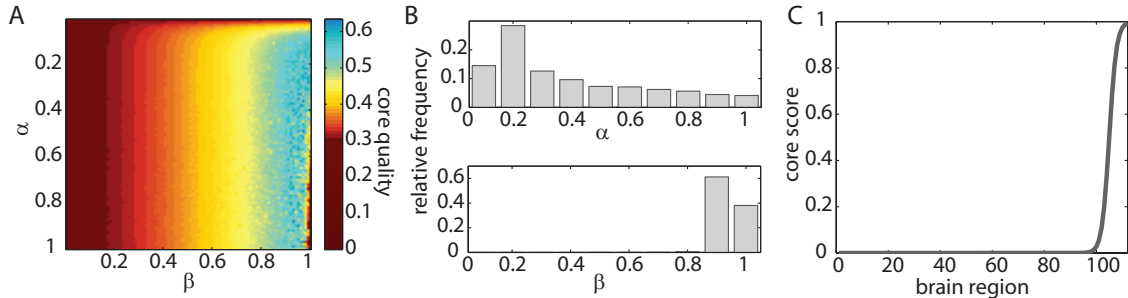


Figure 10: **Geometric Core-Periphery Organisation in Brain Networks.** (A) Core quality  $R(5)$  in the  $(\alpha, \beta)$  parameter plane for a typical participant (3), scanning session (1), sequence type (EXT), and experimental block (1). (B) The distribution of the  $\alpha$  and  $\beta$  values that maximise the  $R$ -Score. We compute this distribution over all network layers, participants, scanning sessions, and sequence types. (C) Mean core shape. Plot of the ordered  $C$  vector with the values of  $\alpha$  and  $\beta$  set to the mean values of those that maximise the  $R$ -score for all network layers, participants, scanning sessions, and sequence types.

network. The roles of nodes in a static network can be studied in multiple ways [44], and we focus on describing their geometric core-periphery organisation to compare it with the temporal core-periphery organisation discussed above.

As described in Ref. [35], the role of a node along a core-periphery spectrum can be examined using a centrality measure known as the (geometric) core score, which we calculate separately for the static networks from each time window. We find that the core-periphery organisation of these static networks is characterised by a single core rather than by multiple cores. In Figure 10A, we show a typical  $R$ -score landscape in the  $(\alpha, \beta)$  parameter plane, which favours a relatively small core (the mean size is set by  $\beta \approx 0.94$ ) and a mid-range value of the transition-sharpness parameter (the mean is  $\alpha \approx 0.40$ ). The former is particularly evident from the Figure 10A.

We now describe in detail our choice of “ideal”  $\alpha$  and  $\beta$  values. In Figure 10B, we show the distributions of the relative frequencies of  $\alpha$  and  $\beta$  values that maximise the  $R$ -score for each network layer, participant, scanning session, and sequence type. The  $\beta$  parameter is much more localised (its standard deviation is 0.05) than the  $\alpha$  parameter (its standard deviation is 0.26).

We now assign a core score to each node using only the mean values of  $\alpha$  and  $\beta$ . In Figure 10C, we show the shape of the ‘mean core’ that we obtain using these parameter values. This figure demonstrates that the typical (geometric) core-periphery organisation in the networks under study is a mixture between a discrete core-periphery organisation, in which every node is either in the core or in the periphery, and a continuous core-periphery organisation, in which there is a continuous spectrum to describe how strongly nodes belong to a core. In these networks (which usually possess a single core), the majority of nodes do not belong to the core, but those nodes that do (roughly 10% of the nodes) have a continuum level of association strengths with the core.

In some cases, we identified multiple competing cores, evidenced by the simulated annealing method exploring a local maximum rather than identifying a global maximum. Because of this stochastically in the algorithm, we ran the analysis with the “ideal”  $\alpha$  and  $\beta$  10 times and used the solution with the highest  $R$  score out of these 10 iterations for each network layer, subject, scanning session, and sequence type.

An interesting question is whether geometric core scores remain relatively constant for a given brain region throughout time or whether they change with learning. We observed that regions that have a high geometric core score in the first scanning session and in EXT blocks

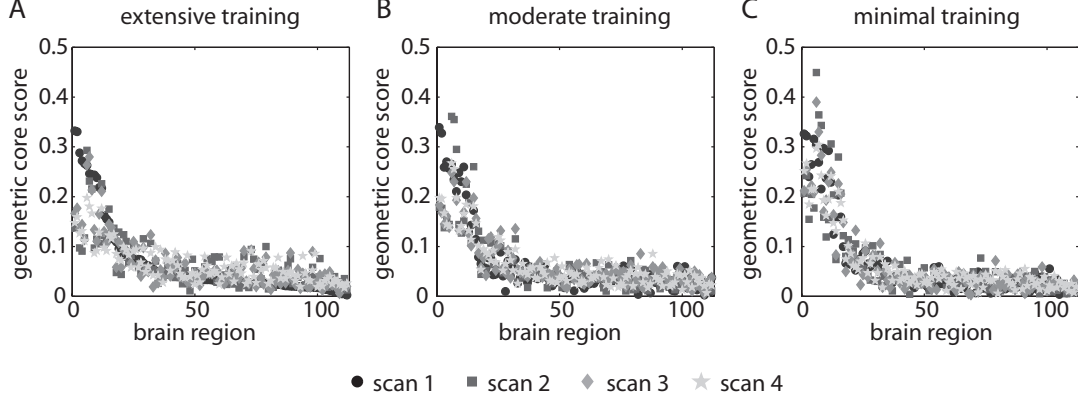


Figure 11: **Geometric Core-Periphery Organisation Over 42 Days.** The geometric core scores for each brain region defined by the trial blocks in which participants practiced sequences that would eventually be (A) extensively trained, (B) moderately trained, and (C) minimally trained for data from scanning sessions 1 (day 1; black circles), 2 (after approximately 2 weeks of training; dark gray squares), 3 (after approximately 4 weeks of training; gray diamonds), and 4 (after approximately 6 weeks of training; light gray stars). We have averaged the geometric core scores over blocks and over 20 participants. The order of brain regions is identical for all 3 panels (A-C), and we chose it by sorting the geometric core scores from the EXT blocks on scanning session 1 (on day 1 of the experiment).

were likely to have high geometric core scores in later scanning sessions and in MOD and MIN blocks; see Figure 11. However, as shown in Figure 12, we also observed that the absolute depth of the core changed as a function of learning. We estimate the absolute depth of the core using the variance of the geometric core scores over brain regions: a high variance indicates a deeper core and a low variance indicates a shallower core. In Figure 12A, we see that the variance of the geometric core-scores over brain regions decreases with learning for all 3 sequence types, and it decreases more drastically for more extended learning (EXT and MOD) than for early learning (MIN). In Figure 12B-C, we observe that this decrease in the variance of geometric core scores over brain regions is associated with both decreased skewness and decreased kurtosis, suggesting that the distribution becomes more symmetric. Together, these results suggest that the geometric core-periphery organisation is altered significantly by learning.

## Relationship Between Temporal and Geometric Core-Periphery Organisation

Given the geometric core-periphery organisation in the individual layers of the multilayer networks and the temporal core-periphery organisation in the full multilayer networks, it is important to ask whether brain regions in the temporal core (i.e., regions with low flexibility) are also likely to be in the geometric core (i.e., whether they exhibit strong connectivity with other core nodes, as represented by a high value of the geometric core score). In Figure 13, we show scatter plots of the flexibility and core score for the 3 training levels (EXT, MOD, and MIN) and the 4 scanning sessions over the 6-week training period. We find that the temporal core-periphery organisation (which is a dynamic measurement) is strongly correlated with the geometric core score (which is a measure of network geometry and hence of network structure). This indicates that regions with low temporal flexibility tend to be strongly-connected core nodes in (static) network layers.

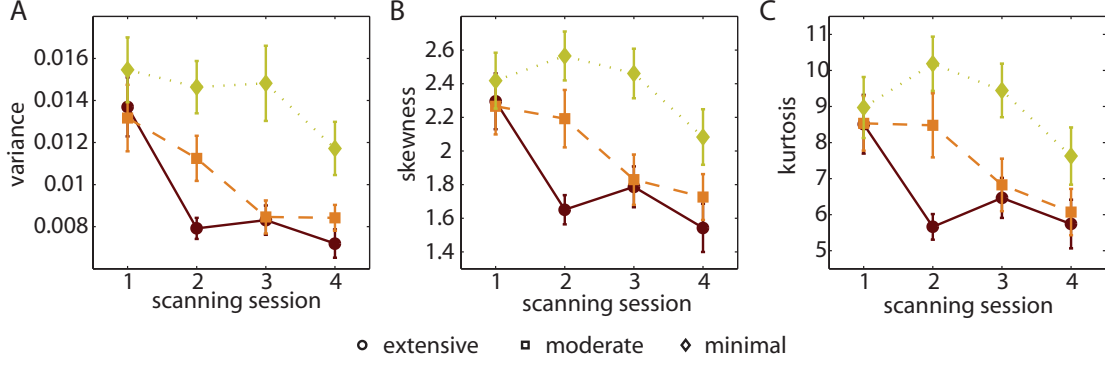


Figure 12: **Geometric Core Scores Change with Learning.** The (A) variance, (B) skewness, and (C) kurtosis of the distribution of geometric core scores over brain regions, which we calculated for blocks of extensively trained (EXT; maroon circles, solid lines), moderately trained (MOD; orange squares, dashed lines), and minimally trained (MIN; gold diamonds, dotted lines) sequences over the 4 scanning sessions spanning the 6 weeks of training. We averaged the values for each moment over blocks and over the 112 brain regions. Error bars indicate the standard error of the mean over participants (where the data point from each participant is the mean geometric core score over brain regions, scanning sessions, sequence types, and network layers).

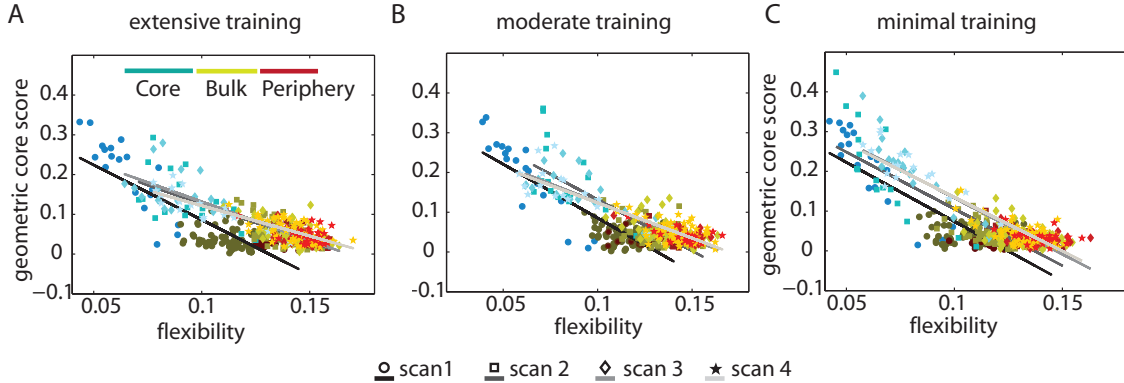


Figure 13: **Relationship Between Temporal and Geometric Core-Periphery Organisations.** for networks constructed from blocks of (A) extensively, (B) moderately, and (C) minimally trained sequences on scanning session 1 (day 1; circles), session 2 (after approximately 2 weeks of training; squares), session 3 (after approximately 4 weeks of training; diamonds), and session 4 (after approximately 6 weeks of training; stars). We show temporal core nodes in cyan, temporal bulk nodes in gold, and temporal periphery nodes in maroon. Colour darkness of data points indicates scanning session with darker colours indicating scan 1 and lighter colours indicating scan 4. Lines with colours ranging from black to light gray indicate the best linear fits in grayscale for session 1 (black) through session 4 (light gray). The Pearson correlation between the flexibility (averaged over 100 multilayer modularity optimisations, 20 participants, and 4 scanning sessions) and the geometric core score (averaged over 20 participants and 4 scanning sessions) is significant for the EXT ( $r \approx -0.91$ ,  $p \approx 3.3 \times 10^{-45}$ ), MOD ( $r \approx -0.92$ ,  $p \approx 2.2 \times 10^{-49}$ ), and MIN ( $r \approx -0.93$ ,  $p \approx 4.7 \times 10^{-50}$ ) data.

## Temporal Core-Periphery Organisation and Network Co-evolution

In our comparison between geometric and temporal core-periphery organisation, we found that nodes in the geometric core of a network tend to have low temporal flexibility (as measured by changes in community allegiance over time). That is, as a function of time, nodes in the geometric core appear to play consistent roles in specific communities. However, it remains unclear whether the edges that emanate from the geometric/temporal core regions of the brain have weights that remain constant (and constantly high) through time or whether these weights instead vary over time in a cohesive manner. Either of these possibilities can yield a strong correlation between geometric and temporal cores. It is therefore important to perform to test these two possibilities against each other. To do this, we exploit the fact that the first possibility (i.e., perpetually strong edge weights emanating from core nodes) suggests a static role for core nodes in cognitive function, whereas the second possibility (i.e., co-varying strong weights of the edges that emanate from core nodes) instead suggests a dynamic role for core nodes in cognitive function.

To determine which pattern best reflects learning-dependent changes of brain dynamics, we draw inspiration from a different biological system: the cellular cytoskeleton [45], in which actin filaments form bridges (edges) between different parts (nodes) of a cell. Importantly, the bridges are themselves linked to one another via actin-binding proteins. Because network edges are not independent, this cross-linking structure has important implications for the mechanical and transport properties of the cytoskeleton. Analogously, one can think of correlations between the time series from different regions of the human brain as information bridges like the actin filaments. Moreover, one can think of the time-dependent relationships between these correlations as cross-links that change the temporal landscape for information processing.

To describe the cross-linking structure in dynamic human brain networks, we quantify the co-variability of edge weights through time using a hypergraph formalism. Each of the hyperedges consists of the set of edges that co-vary significantly in weight with one another over time, and the size (i.e., cardinality)  $s$  of a hyperedge is equal to the number of edges it contains. The set of all hyperedges identified across the temporal networks forms a hypergraph.

We first study an *early-learning* hypergraph, which is associated with the temporal networks of scanning session 1 (including EXT, MOD, and MIN blocks). The set of networks that yield this hypergraph includes edge-weight time series whose length is approximately 30 time points (i.e., 10 each of extensively, moderately, and minimally trained blocks; see Table 2). The distribution of hyperedge sizes is heavy-tailed (see Figure 14A). We evaluated the size distribution using a statistical approach developed by Clauset et al. [46] that combines maximum-likelihood fitting methods with goodness-of-fit tests based on the Kolmogorov-Smirnov statistic and likelihood ratios. This computation confirmed that this distribution is well-fit by a power law [47] at the 0.1 significance level: the p-value was  $p \approx 0.875$ , which is well above the threshold of  $p = 0.1$  for ruling out a power law. Hence, the distribution is very well fit by  $P_s \sim s^\nu$  with exponent  $\nu \approx -2.69$ , an estimated lower bound for the onset of power-law behaviour equal to  $s_{\min} \approx 8$ , a log-likelihood of the data  $s \geq s_{\min}$  under the fitted power law of  $L \approx -274.45$ , and a Kolmogorov-Smirnov goodness-of-fit statistic of  $D \approx 0.0364$ .

To identify how hyperedges are distributed in the brain’s anatomical space, we calculated the hypergraph node degree of each brain region. (We define the hypergraph *degree* of a node as the number of hyperedges incident to the node.) In Figure 14B, we illustrate that nodes with high hypergraph degree are located predominantly in the primary sensorimotor cortex. Observe that this distribution among brain regions has striking similarities to the temporal core shown in Figure 2 in the main manuscript. We quantify this similarity by computing the Pearson correlation coefficient between the temporal core-periphery organisation (as measured using flexibility) and the structure of network co-variability (as measured using hypergraph node degree), and we show the result as a scatter plot in Figure 14C. The strong negative correlation between these two diagnostics indicates that the temporal core is composed predominantly of

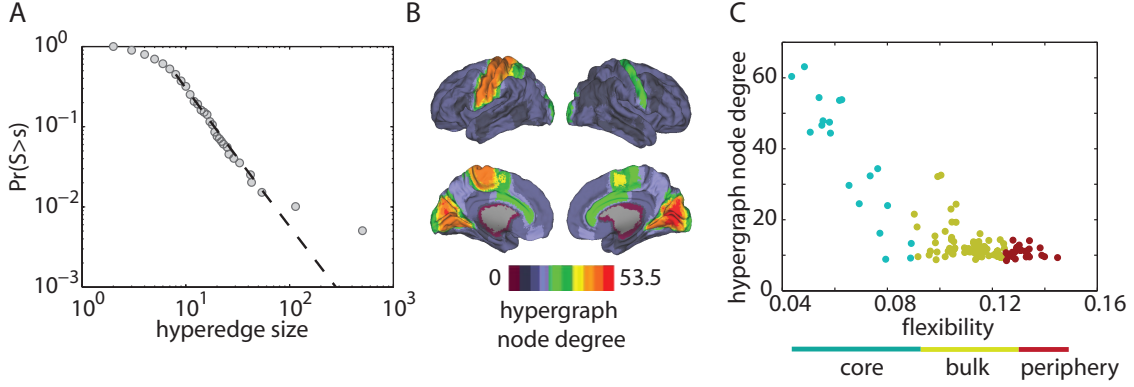


Figure 14: **Structure of Early-Learning Hypergraphs.** (A) Probability distribution of the size  $s$  (i.e., cardinality) of hyperedges. The fitted line indicates the best power-law fit, which we identified using the statistical approach of Clauset et al. [46]. We use a p-value of greater than 0.1 to indicate the identification of significant power-law behaviour. (B) Anatomical distribution of hypergraph node degree (averaged over participants). Nodes with high hypergraph degree are evident in primary sensorimotor regions. (C) Scatter plot of the early-learning hypergraph node degree (averaged over the 20 participants) and flexibility (averaged over the 100 multilayer modularity optimisations and the 20 participants) of the scan-1 EXT multilayer networks. We computed a Pearson correlation coefficient of  $r \approx -0.82$  with a p-value of  $p \approx 6.8 \times 10^{-29}$ . We show the brain regions from the temporal core in cyan, the regions from the temporal bulk in gold, and the regions from the temporal bulk in maroon. Regions in the temporal core tend to participate in the largest number of hyperedges.

nodes with high hypergraph degree. That is, regions with low temporal flexibility in community allegiance tend to have connections with other nodes with which they co-vary strongly in weight over time.

To examine later learning, we constructed 3 families of *extended-learning hypergraphs*: the first (which we call an ‘EXT hypergraph’) is composed of the EXT blocks in scanning sessions 2–4, the second (a ‘MOD hypergraph’) is composed of the MOD blocks in scanning sessions 2–4, and the third (a ‘MIN hypergraph’) is composed of the MIN blocks in scanning sessions 2–4. We find that all 3 families of extended-learning hypergraphs also displayed heavy-tailed hyperedge size distributions, although only the MIN hypergraph can be construed to be well-fit by a power-law at the 0.1 significance level (see Figure 15A). Furthermore, consistent with our results on early-learning hypergraphs, we find that regions in the temporal core tend to participate in a larger number of extended-learning hyperedges than regions in the temporal bulk or temporal periphery. The strong negative correlation between flexibility and hypergraph node degree again suggests that regions with low temporal flexibility in community allegiance tend to have connections with other nodes with which they co-vary strongly in weight over time.

To gain intuition for how these hyperedges might affect brain communication, we examine which types of nodes the hyperedges tend to connect. For example, do the hyperedges tend to connect core nodes with other core nodes, do they tend to connect core nodes with periphery nodes, or are both situations common. Recall that  $\xi$  denotes the adjacency matrix for the edge-edge co-variation network and that the set of connected components (hyperedges) in this network forms a hypergraph. We define a *link* between two nodes (i.e., brain regions) in the hypergraph to exist if those two nodes participate in the same hyperedge. To study the co-participation of nodes in hyperedges, we construct the link matrix  $\varrho$  to be the  $N \times N$  adjacency matrix with elements  $\varrho_{ij}$  that are equal to 1 if nodes  $i$  and  $j$  participate in the same hyperedge



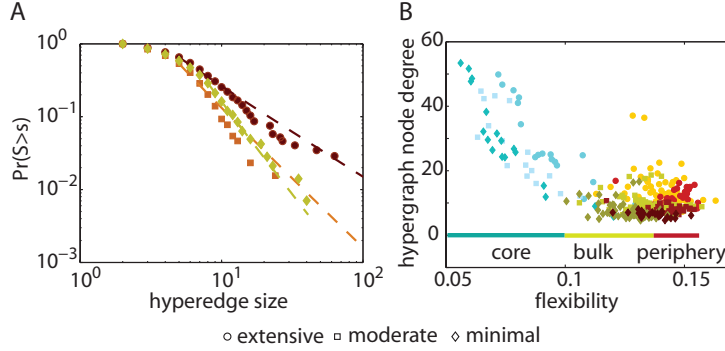


Figure 15: **Structure of Extended-Learning Hypergraphs.** (A) Probability distribution of the size  $s$  of hyperedges for the EXT (maroon; circles), MOD (orange; squares), and MIN (gold; diamonds) hypergraphs. The fitted lines indicate the best power-law fit, which we identified using the statistical approach of Clauset et al. [46]: the Kolmogorov-Smirnov goodness-of-fit statistics were  $D \approx 0.0767$  for EXT hypergraphs,  $D \approx 0.0845$  for MOD hypergraphs, and  $D \approx 0.0383$  for MIN hypergraphs. The power-law exponents were generally high ( $v \approx -2.22$  for EXT,  $v \approx -2.88$  for MOD, and  $v \approx -3.47$  for MIN), and the p-values for the acceptance of the null hypothesis indicated that only the size distribution of the MIN hypergraph could be said to be well-fit by a power-law at the 0.1 significance level:  $p \approx 0.01$  for EXT,  $p \approx 0.03$  for MOD, and  $p \approx 0.95$  for MIN. The estimated lower bounds of the power-law behaviour are small ( $s_{\min} \approx 5$  for both EXT and MOD hypergraphs, and  $s_{\min} \approx 8$  for MIN hypergraphs), and the log-likelihood of the data for  $s \geq s_{\min}$  under the fitted power law are  $L \approx -353.94$  for EXT hypergraphs,  $L \approx -166.86$  for MOD hypergraphs, and  $L \approx -103.46$  for MIN hypergraphs. (B) Scatter plots of the EXT, MOD, and MIN hypergraph node degrees (averaged over the 20 participants) versus the flexibility averaged, respectively, over the EXT multilayer networks from scans 2–4 (circles), the MOD multilayer networks from scans 2–4 (squares), and the MIN multilayer networks from scans 2–4 (diamonds). The corresponding Pearson correlation coefficients and p-values are  $r \approx -0.77$  and  $p \approx 9.0 \times 10^{-24}$  for EXT hypergraphs,  $r \approx -0.84$  and  $p \approx 9.2 \times 10^{-32}$  for MOD hypergraphs, and  $r \approx -0.88$  and  $p \approx 2.0 \times 10^{-38}$  for MIN hypergraphs. Note that we also averaged flexibility values over the 100 multilayer modularity optimisations and the 20 participants. We again show the temporal core regions in cyan, the temporal bulk regions in gold, and the temporal periphery regions in maroon. Colour darkness of data points indicates scanning session with darker colours indicating scan 1 and lighter colours indicating scan 4. Consistent with our results on early learning, brain regions in the temporal core tend to participate in the largest number of extended-learning hyperedges.



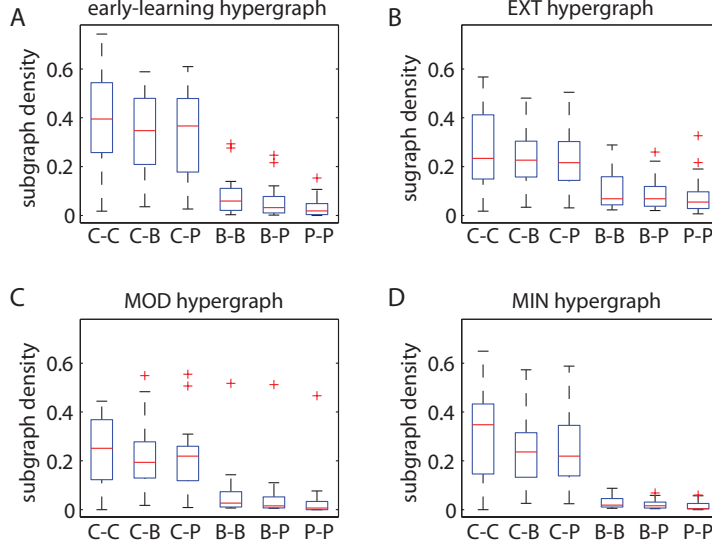


Figure 16: **Hyperedges Among Core, Bulk, and Periphery Regions** Network density for the core-core ( $C-C$ ) subgraph  $\varrho_{C-C}$ , the core-bulk ( $C-B$ ) subgraph  $\varrho_{C-B}$ , the core-periphery ( $C-P$ ) subgraph  $\varrho_{C-P}$ , the bulk-bulk ( $B-B$ ) subgraph  $\varrho_{B-B}$ , the bulk-periphery ( $B-P$ ) subgraph  $\varrho_{B-P}$ , and the periphery-periphery ( $P-P$ ) subgraph  $\varrho_{P-P}$ . Hyperedges appear to emanate preferentially from core nodes; hence, they tend to connect core nodes with other core nodes, with nodes in the bulk, and with nodes in the periphery.

and are equal to 0 otherwise. We quantify the probability of a hyperedge linking two core nodes using the network density of the *core-core subgraph*  $\varrho_{C-C}$  of the link matrix  $\varrho$ . (The subscript  $C - C$  indicates that we only keep elements in the adjacency matrix that connect a core node with another core node.) We similarly define the core-bulk subgraph  $\varrho_{C-B}$ , the core-periphery subgraph  $\varrho_{C-P}$ , the bulk-bulk subgraph  $\varrho_{B-B}$ , the bulk-periphery subgraph  $\varrho_{B-P}$ , and the periphery-periphery subgraph  $\varrho_{P-P}$ . The network density of a subgraph is defined as the number of non-zero entries in the subgraph divided by the total number of possible non-zero entries. By computing these densities, we find that hyperedges tended to link core nodes with other core nodes, with nodes in the bulk, and with nodes in the periphery but do not tend to link nodes in the bulk with other nodes in the bulk, nodes in the bulk with nodes in the periphery, or nodes in the periphery with other nodes in the periphery (see Figure 16). This indicates that hyperedges emanate preferentially from core nodes.

## Methodological Considerations

### Experimental Factors

#### Effect of Region Size

Recent studies have noted an effect of region size on estimates of hard-wired connectivity strength used in constructing structural connectomes [48, 49]. Although the present work is concerned with functional connectomes, it is nevertheless relevant to consider whether or not region size could be a driving effect of the observed core-periphery organisation. Importantly, we observe no significant correlation between region size and flexibility, suggesting that region size is not driving the reported results (see Figure 17).

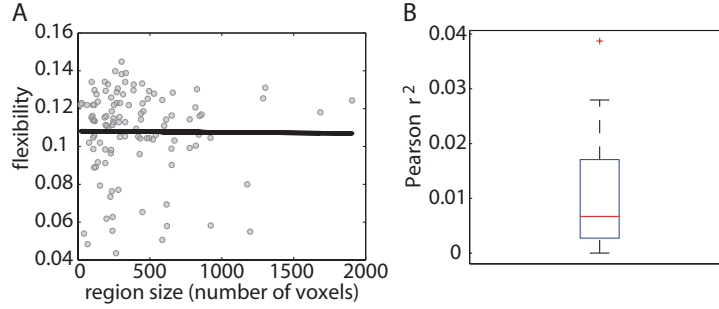


Figure 17: **Region Size is Uncorrelated with Flexibility.** (A) Scatter plot of the size of the brain region in voxels (averaged over participants) versus the flexibility of the EXT multilayer networks, which we averaged over the 100 multilayer modularity optimisations and the 20 participants. Data points indicate brain regions. The line indicates the best linear fit. Its Pearson correlation coefficient is  $r \approx -0.009$ , and the associated p-value is  $p \approx 0.92$ . (B) Box plot over the 20 participants of the squared Pearson correlation coefficient  $r^2$  between the participant-specific region size in voxels and the participant-specific flexibility averaged over the 100 multilayer modularity optimisations.

### Effect of Block Design

Another important factor is the underlying experimental block design and its effect on the correlation structure between brain regions within a single time window (i.e., within a single layer in the multilayer formalism). Two brain regions (e.g., M1 and SMA) might be active during the trial but quiet during the inter-trial interval (ITI), leading to a characteristic on-off activity pattern that is highly correlated with all other regions that also turn on with the task and off during the ITI. The frequency of this task-based activity (one on-off cycle per trial, where each trial is of length 4–6 TRs) is included in our frequency band of interest (wavelet scale two, whose frequency range is 0.06–0.12 Hz), and it therefore likely plays a role in the observed correlation patterns between brain regions in a single time window.

We note, however, that our investigations of dynamic network structure—namely, our computations of flexibility of community allegiance and edge-edge correlations over multiple time windows—probe functional connectivity dynamics at much larger time scales, and the associated frequencies are an order of magnitude smaller. They lie in the range 0.0083–0.012 Hz, as there is one time window every 40–60 TRs. At these longer time scales, we can probe the effects of both early learning and extended learning independent of block-design effects.

### Specificity of Dynamic Network Organisation as a Predictor of Learning

An important consideration is whether there exist (arguably) simpler properties of brain function than flexibility that could be used to predict learning. We find that the power of activity, the mean connectivity strength, and parameter estimates from a general linear model provide less predictive power than flexibility.

**Measures of Activity and Connectivity** Although it is not within the scope of this study to perform an exhaustive search of all possible measures of brain-region activity, we focus on two commonly used diagnostics. One is based on functional connectivity, and the other is based on brain activity. To estimate the strength of functional connectivity, we calculated the mean pairwise coherence between regional wavelet scale two time series constructed from the BOLD signal, where the mean was taken over all possible pairs of regions and all EXT experimental blocks extracted from scans on day 1 for a given subject. To estimate the strength of activity,

we calculated the mean signal power of the regional wavelet scale two time series constructed from the BOLD signal, where the mean was taken over all regions and all EXT experimental blocks extracted from scans on day 1 for a given subject. We estimate the power of the wavelet two time series  $P_{w_2}$  as the square of the time series normalised by its length:  $P_{w_2} = \sum_t \frac{w_2(t)^2}{T}$  where  $T$  is the length of the time series [50, 51]. We found that neither mean pairwise coherence nor mean power of regional activity could be used to predict learning during the following 10 home training sessions. For the mean pairwise coherence, we obtained a Pearson correlation of  $r \approx -0.003$  and a p-value of  $p \approx 0.987$ . For the mean power of brain-region activity, we obtained  $r \approx -0.218$  and  $p \approx 0.354$ . These results indicate that a prediction similar to that made using flexibility is not possible using the (arguably) simpler properties of the mean pairwise coherence or the mean power of regional brain activity. Moreover, they also suggest that the dynamic pattern of coherent functional brain activity is more predictive than the mean.

**Parameter Estimates for General Linear Model** We determined relative differences in the BOLD signal by using a general linear model (GLM) approach for event-related functional data [52]. For each participant, we constructed a single design matrix for event-related fMRI by specifying the onset time and duration of all stimulus events from each scanning session (i.e., the pre-training session and the 3 test sessions). We found estimations of BOLD-signal related change for conditions of interest by using the design matrix with the General Linear Model (GLM). We modelled the duration of each sequence trial as the time elapsed to produce the entire sequence, or movement time (MT). MT is a direct measure of the time spent on task, and leads to accurate modelling of BOLD using the GLM [53]. Separate stimulus vectors indicate each sequence exposure type (EXT, MOD, and MIN) separately for each scanning session. We took potential differences in brain activity due to rate of movement into account by using the MT for each trial as the modelled duration for the corresponding event. We convolved events using the canonical hemodynamic response function and temporal derivative. Using freely available software [54], we then combined the corresponding beta image pairs for each event type (HRF and temporal derivative) at the voxel level to form a magnitude image [55]

$$H = \text{sign}(\hat{B}_1) + \sqrt{(\hat{B}_1 + \hat{B}_2)}, \quad (8)$$

where  $H$  is called the ‘combined amplitude’ of the estimation of BOLD ( $\hat{B}_1$ ) and its temporal derivative ( $\hat{B}_2$ ).<sup>1</sup> This yielded separate magnitude images for each sequence exposure type (EXT, MOD, and MIN) and session. We then calculated the mean region-based magnitude for each exposure type and session using those regions derived from each subject’s grey matter-constrained HO atlas. This was restricted to those voxels with positive values within the magnitude images, which reflect BOLD signal change that is above a resting baseline.

We did not find a significant correlation between the mean parameter estimates averaged over brain regions for the EXT trials in scanning session 1 and learning of the EXT sequences over the subsequent approximately 10 home training sessions. The Pearson correlation is  $r \approx -0.10$  and the p-value is  $p \approx 0.65$ ; compare these results to those shown in Figure 4.

## Dynamic Community Detection

In the multilayer modularity quality function (2), we need to choose values for two parameters [41]: a structural resolution parameter  $\gamma$  and a temporal resolution parameter  $\omega$ . In the following two sections, we examine the effects of these choices on our results.

---

<sup>1</sup>In this equation, we use the hat notation to indicate the sample estimate from a population parameter value (i.e., a single participant).

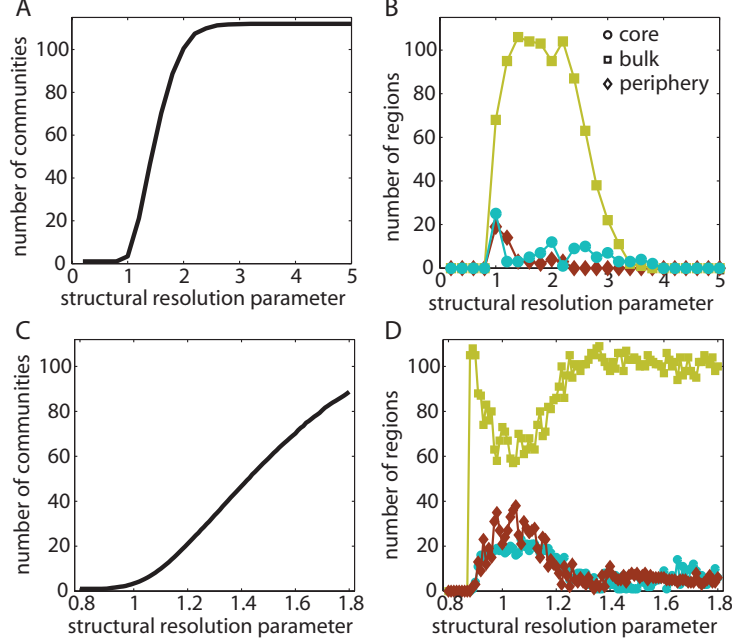


Figure 18: **Effect of Structural Resolution Parameter.** (A,C) Number of communities and (B,D) number of regions in the temporal core (cyan; circles), temporal bulk (gold; squares), and temporal periphery (maroon; diamonds) as a function of the structural resolution parameter  $\gamma$ , where we considered (A,B)  $\gamma \in [0.2, 5]$  in increments of  $\Delta\gamma = 0.2$  and (C,D)  $\gamma \in [0.8, 1.8]$  in increments of  $\Delta\gamma = 0.01$ . We average the values in panels (A) and (C) over 100 multilayer modularity optimisations and over the 20 participants.

### Effect of Structural Resolution Parameter

For the results that we report in the main manuscript, we have chosen to use the value  $\gamma = 1$ , which is the most common choice in many community detection studies in the optimisation of the single-layer and multilayer modularity quality function [28, 29, 56]. In this case, one compares the adjacency tensors of the real data  $\mathbf{A}$  to the adjacency matrices of the optimisation null model  $\mathbf{P}$  via subtraction; that is, one examines the tensor  $\mathbf{A} - \mathbf{P}$ . However, one can lower  $\gamma$  to access community structure at smaller spatial scales (i.e., to examine smaller communities) or increase it to access community structure at larger spatial scales (i.e., to examine larger communities). By examining network diagnostics over a range of  $\gamma$  values, we are able to investigate the spatial specificity of our results.

We observe that the mean number of communities in the partitions obtained by optimising multilayer modularity (2) varies from the minimum (1) to the maximum (112) possible values for  $\gamma$  approximately in the interval  $[0.8, 2.5]$  (see Figure 18A). We investigate this transition in greater detail in Figure 18C-D. Near the value  $\gamma = 1$ , the number of regions in the bulk dips to about 65, whereas the number of regions in the core and periphery rise to about 20 and 25, respectively. This combined dip-rise feature extends over approximately the  $\gamma$  range  $(0.8, 1.2)$ , which corresponds to partitions that are composed of between approximately 3 and approximately 15 communities (with an associated mean community size of between approximately 8 and approximately 37 brain regions). This supports our claim that the temporal core-periphery structure that we examine in this study is a genuine meso-scale feature of coherent brain dynamics.

## Effect of Temporal Resolution Parameter

For the results that we report in the main manuscript, we used  $\omega = 1$ . The value  $\omega = 1$  ensures that the *inter*-layer coupling is equal to the maximum value of the *intra*-layer coupling, which we compute from the magnitude squared coherence (which is constrained to lie in the interval  $[0, 1]$ ). This gives some support for the temporal resolution parameter value  $\omega = 1$ , but it is important to check the robustness of results for different values of this parameter and investigating dynamic network structure at other values of  $\omega$  can also provide additional insights [41]. For example, one can decrease  $\omega$  to allow greater variability in the partitions of nodes across individual layers (i.e., across time in temporal networks) or increase it to constrain the partitions to be more similar across layers (i.e., across time in temporal networks). By examining network diagnostics over a range of  $\omega$  values, we can also quantify the robustness of our results to differing amounts of temporal variation in community structure.

We varied  $\omega$  from 0.1 to 2 in increments of  $\Delta\omega = 0.1$ . As expected, we find that the number of communities identified in the optimisation of the multilayer modularity quality function decreases as  $\omega$  is increased (see Figure 19A). This is consistent with the fact that greater variation of partitions across time is possible for smaller values of  $\omega$ . Variation between partition of nodes in individual layers can occur in two ways: (1) a small number of regions change community membership from one layer to the next, but the majority of regions retain their community membership; or (2) entire communities fragment such that the algorithm identifies either the “death” of a community that was present in the previous layer but is not present in the current layer or the birth of a “new” community that was not present in the previous layer but is present in the current layer. The second type of behaviour leads to an increased number of different communities over time.

For each value of *omega*, we examined the robustness of our delineation of brain regions into a temporal core, a temporal bulk, and a temporal periphery using the same procedure that we employed for  $\omega = 1$ . Namely, we defined a temporal core and temporal periphery as those brain regions that were composed, respectively, of the brain regions below and above the 95% confidence interval of the nodal null model. We report the number of regions in each group as a function of  $\omega$  in Figure 19B. Strikingly, the number of brain regions that we identified as being part of the temporal core varied little over the examined range of  $\omega$  values; it remained at approximately  $17.0 \pm 1.1$ . The number of regions in the temporal bulk and temporal periphery varied more (with values of approximately  $75.6 \pm 7.4$  for the bulk and approximately  $19.4 \pm 6.8$  for the periphery), suggesting that the separation between the temporal bulk and temporal periphery is less drastic than that between temporal core and temporal bulk. Indeed, this difference in separation of core versus bulk and bulk versus periphery is evident in Figure 2 of the main manuscript and in Figures 7 and 8 of this supplement.

Finally, we tested the robustness of our predictions of learning to variations in the temporal resolution parameter  $\omega$ . We find that we can predict learning in the first 10 days of home training using the flexibility in the EXT blocks of scan 1 estimated at  $\omega$  values between 0.8 and 1.5 (see Figure 19C). Indeed, it is important to note that in reporting our results in the main manuscript, we very specifically did not cherry-pick an  $\omega$  value that maximises our predictive sensitivity (which occurs at  $\omega \approx 1.4$  and  $\omega \approx 0.8$ ). Instead, we report results at  $\omega = 1$  that are representative of results that we obtain over a wide range of  $\omega$  values.

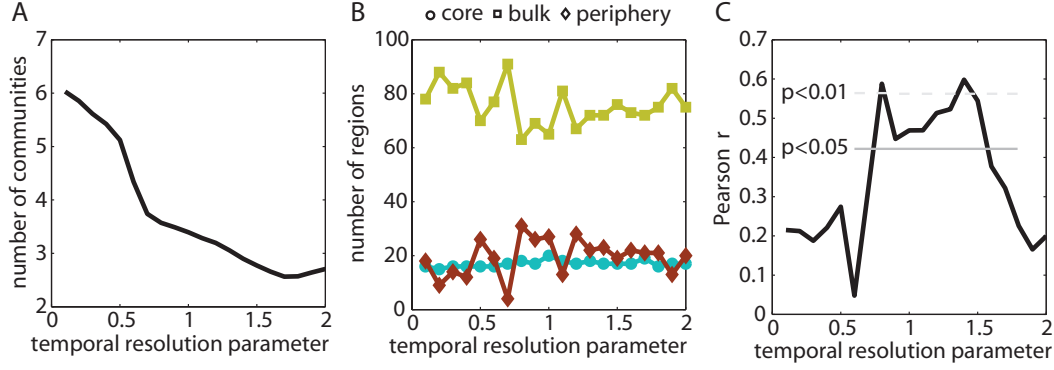


Figure 19: **Effect of Temporal Resolution Parameter.** (A) The number of communities averaged over 100 multilayer modularity optimisations and over 20 participants as a function of the temporal resolution parameter  $\omega$ . (B) The number of regions that we identified as part of the temporal core (cyan; circles), temporal bulk (gold; squares), and temporal periphery (maroon; diamonds) as we vary  $\omega$  from 0.1 to 2 in increments of  $\Delta\omega = 0.1$ . (C) Pearson correlation coefficient  $r$  between the flexibility (averaged over 112 brain regions and 100 multilayer modularity optimisations) from the EXT blocks in scan 1 and learning over the subsequent 10 days of home training (which we estimated using the learning parameter  $\kappa$ ) as a function of  $\omega$ . The dark gray (solid) horizontal line indicates the threshold  $r$  value above which correlations are significant at a p-value of  $p < 0.05$ ; the lighter gray (dashed) horizontal line indicates the threshold  $r$  value above which correlations are significant at  $p < 0.01$ . For comparison, see the scatter plot of  $\kappa$  versus the mean flexibility at  $\omega = 1$  in Figure 4.

## Supplementary Discussion

### Network Predictions of Future Learning

In this study, we have observed that properties of the temporal organisation of functional brain networks (e.g., on day 1 of this experiment) can be used to predict learning (e.g., on the following 10 days of home training in this experiment). Our findings are consistent with two previous studies that demonstrated a predictive link between both dynamic [26] and topological [57] network organisation and subsequent learning. In the first study, dynamic network flexibility on the first day of a simple motor learning task (cued sequence production) predicted learning on the second day, and flexibility on the second day predicted learning on the third day [26]. The second study investigated participants' success in learning words of an artificial spoken language and found that network properties from individual time windows could be used to predict this success [57]. Importantly, together with the present study, these results highlight the potential breadth of the relationship between network organisation and learning. The presence of such a relationship has now been identified across multiple tasks, over multiple time scales, and using both dynamic and topological network properties.

### Dynamic Brain Networks

The fact that functional connectivity in the brain changes over time, and that these changes are biologically meaningful, is becoming increasingly apparent. Several recent studies have highlighted the temporal variability [58, 59, 60] and non-stationarity [61] of functional brain network organisation, which are both more apparent over shorter time intervals and less apparent over longer time intervals [58, 60, 59]. Although temporal variability in functional connectivity, was

seen initially as a signature of measurement noise [61], recent evidence suggests instead that it might provide an indirect measurement of changing cognitive processes and hence serve as a diagnostic biomarker of disease [61, 62]. Moreover, such temporal variability appears to be modulated by exogenous inputs. For example, Barnes et al. [63] have demonstrated using a continuous acquisition ‘rest-task-rest’ design that endogenous brain dynamics do not return to their pre-task state until approximately 18 minutes following task completion. Similar results that consider other tasks have also been reported [64]. More generally, the dynamic nature of brain connectivity is likely linked to spontaneous cortical processing, reflecting a combination of both stable and transient communication pathways [60, 59].

## Modularity Versus Core-Periphery Organisation

Modularity and core-periphery organisation are two types of meso-scale structures, and they can both be present simultaneously in a network [35, 65]. Moreover, the two types of organisation can in principle pertain to different characteristics of or constraints on underlying brain function. In particular, the presence of community structure supports the idea of the brain containing putative functional modules, whereas the presence of a core-periphery organisation underscores the fact that different brain regions likely play inherently different roles in information processing. A symbiosis between these two types of organisation is highlighted by the findings that we report in this manuscript: the dynamic reconfiguration of putative functional modules can be described parsimoniously by temporal core-periphery organisation, demonstrating that one type of meso-scale structure can help to characterise another. Furthermore, the notion that the brain can simultaneously contain functional modules (e.g., the executive network or the default-mode network) and regions that transiently mediate interactions between modules is consistent with recent characterisations of attention and cognitive control processes [66].

## Core-Periphery Organisation of Human Brain Structure and Function

The notion of a core-periphery organisation is traditionally based on the structure (rather than the temporal dynamics) of a network. In this paper, we have described this using the terminology *geometric core-periphery organisation*. (It is geometric rather than topological because the networks are weighted.) This notion was formalised in social networks by Borgatti and Everett in 1999 [67]. Available methods to identify and quantify geometric core-periphery organisation in networks include ones based on block models [67],  $k$ -core organisation [68], and aggregation of information about connectivity and short paths through a network [69]. In general, however, these methods to study geometric core-periphery organisation have required one to binarise networks that are inherently weighted, which in turn requires one to throw away a lot of important information. Moreover, even the recently developed weighted extensions of the  $k$ -core decomposition [70] require a discretisation of  $k$ -shells. Importantly,  $k$ -core decomposition is based on a strong and specific type of core connectivity, so this measure will miss other important core-like structures [35, 65]. A well-known measure called the ‘rich-club coefficient’ (RCC) [71], considers a different but somewhat related question of whether nodes of high degree tend to connect to other nodes of high degree. (The RCC is therefore a form of assortativity.) The RCC has also been extended to weighted networks [72], but it still requires one to specify  $k$ -shells.

The aforementioned limitations notwithstanding, several of the measures discussed above have recently been used successfully to identify a structural core of the human brain white-matter tract network, which is characterised not only by a high  $k$ -core [48] and associated RCC [73, 74] but also by a knotty centre of nodes that have a high geodesic betweenness centrality but not necessarily a high degree [65]. A  $k$ -core decomposition has also been applied to functional brain imaging data to demonstrate that core centrality is higher during correct preparation trials

than during incorrect preparation trials, suggesting altered preparatory adaptations necessary for successful task performance [75].

A novel approach that has been able to overcome many of these limitations is the geometric core-score [35], which is an inherently continuous measure, is defined for weighted networks, and can be used to identify regions of a network core without relying solely on their degree or strength (i.e., weighted degree). Moreover, by using a continuous measure, we can produce (1) continuous results, which allow us to measure whether a brain region is more core-like or periphery-like; (2) a discrete classification of core versus periphery; or (3) a finer discrete division (e.g., into 3 or more groups). In addition, this method can identify multiple geometric cores in a network and rank nodes in terms of how strongly they participate in different possible cores. This sensitivity is particularly helpful for the examination examining brain networks, because (for example) multiple cores are hypothesised to mediate multimodal integration [76]. By using this method, we have been able to demonstrate that functional brain networks derived from task-based data acquired during goal-directed brain activity show geometric core-periphery organisation. Moreover, are specifically characterised by a straightforward core-periphery landscape that includes a relatively small core composed of roughly 10% or so of the nodes in the network.

Importantly, we have introduced in this paper a method and associated definitions to identify a *temporal* core-periphery organisation based on dynamic changes in a node’s module allegiance over time. Our approach is inspired by the notion that although the brain uses the function of a some subset of regions to perform a given task, a set of additionally regions that are associated more peripherally with the task might also be activated in a transient manner. Indeed, several recent studies have highlighted the possibility of a separation between groups of regions that are consistently versus transiently activated during task-related function [77, 78], and they have demonstrated that correlations between such regions can be altered depending on their activity [77, 79]. In the present paper, we demonstrate that the cortex organisation of the entire brain network can be decomposed into a temporal core, whose regions consistently display coherent activity with the same putative functional modules over time and a temporal bulk/periphery, whose regions vary in their module affiliation through time. Moreover, the dynamic network representation of coherent activity between brain regions enables us to quantitatively characterise the temporal core-periphery structure and examine its properties in the context of motor learning.

**Acknowledgments.** We thank Jean M. Carlson, Matthew Cieslak, John Doyle, Daniel Greenfield, and Megan T. Valentine for helpful discussions; John Bushnell for technical support; and James Fowler, Sang Hoon Lee, and Melissa Lever for comments on earlier versions of the manuscript. D.S.B. was supported by the Sage Center for the Study of the Mind, the Errett Fisher Foundation, the Templeton Foundation, the David and Lucile Packard Foundation, the Public Health Service Grant NS44393, and the Institute for Collaborative Biotechnologies through Contract W911NF-09-D-0001 from the US Army Research Office. M.A.P. and M.P.R. acknowledge a research award (#220020177) from the James S. McDonnell Foundation. M.A.P. also acknowledges a grant from the EPSRC (EP/J001759/1). P.J.M. acknowledges support from Award Number R21GM099493 from the National Institute of General Medical Sciences. The content is solely the responsibility of the authors and does not necessarily represent the official views of any of the funding agencies.



	Mean	Minimum	Maximum	Standard Error
Extensively Trained Blocks				
During Scan 1	6.00	9.70	10.00	0.21
During Scan 2	10.00	10.00	10.00	0.00
During Scan 3	10.00	10.00	10.00	0.00
During Scan 4	8.00	9.90	10.00	0.10
Moderately Trained Blocks				
During Scan 1	5.00	9.70	11.00	0.27
During Scan 2	10.00	10.00	10.00	0.00
During Scan 3	10.00	10.00	10.00	0.00
During Scan 4	8.00	9.90	10.00	0.10
Minimally Trained Blocks				
During Scan 1	7.00	9.80	11.00	0.18
During Scan 2	10.00	10.00	10.00	0.00
During Scan 3	10.00	10.00	10.00	0.00
During Scan 4	8.00	9.90	10.00	0.10
Length of Extensively Trained Blocks				
During Scan 1	52.50	61.94	72.20	1.34
During Scan 2	35.50	42.36	45.90	0.72
During Scan 3	35.40	40.79	45.50	0.77
During Scan 4	34.60	40.30	45.70	0.87
Length of Moderately Trained Blocks				
During Scan 1	50.80	61.67	72.60	1.26
During Scan 2	39.70	47.56	57.20	0.80
During Scan 3	37.60	45.07	52.80	0.67
During Scan 4	37.60	43.83	50.60	0.79
Length of Minimally Trained Blocks				
During Scan 1	52.10	61.19	70.60	1.29
During Scan 2	44.10	50.02	57.70	0.73
During Scan 3	42.50	47.37	54.50	0.71
During Scan 4	39.70	45.79	54.10	0.70

Table 2: **Experimental Details for Brain Imaging Data Acquired During Scanning Sessions.** In the top three rows, we give the mean, minimum, maximum, and standard error over participants for the number of blocks composed of extensively, moderately, and minimally trained sequences during scanning sessions. In the bottom three rows, we give (in TRs) the mean, minimum, maximum, and standard error of the length over blocks composed of extensively, moderately, and minimally trained sequences during scanning sessions.

Frontal pole	Cingulate gyrus, anterior
Insular cortex	Cingulate gyrus, posterior
Superior frontal gyrus	Precuneus cortex
Middle frontal gyrus	Cuneus cortex
Inferior frontal gyrus, pars triangularis	Orbital frontal cortex
Inferior frontal gyrus, pars opercularis	Parahippocampal gyrus, anterior
Precentral gyrus	Parahippocampal gyrus, posterior
Temporal pole	Lingual gyrus
Superior temporal gyrus, anterior	Temporal fusiform cortex, anterior
Superior temporal gyrus, posterior	Temporal fusiform cortex, posterior
Middle temporal gyrus, anterior	Temporal occipital fusiform cortex
Middle temporal gyrus, posterior	Occipital fusiform gyrus
Middle temporal gyrus, temporooccipital	Frontal operculum cortex
Inferior temporal gyrus, anterior	Central opercular cortex
Inferior temporal gyrus, posterior	Parietal operculum cortex
Inferior temporal gyrus, temporooccipital	Planum polare
Postcentral gyrus	Heschl's gyrus
Superior parietal lobule	Planum temporale
Supramarginal gyrus, anterior	Supercalcarine cortex
Supramarginal gyrus, posterior	Occipital pole
Angular gyrus	Caudate
Lateral occipital cortex, superior	Putamen
Lateral occipital cortex, inferior	Globus pallidus
Intracalcarine cortex	Thalamus
Frontal medial cortex	Nucleus Accumbens
Supplemental motor area	Parahippocampal gyrus
Subcallosal cortex	Hippocampus
Paracingulate gyrus	Brainstem

Table 3: **Brain regions in the Harvard-Oxford Cortical and Subcortical Parcellation Scheme provided by FSL [13, 14].**

# Bibliography

- [1] Matsumoto, M. & Nishimura, T. Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator. *ACM Transactions on Modeling and Computer Simulation* **8**, 3–30 (1998).
- [2] Schmidt, R. A. & Lee, T. D. *Motor Control and Learning A Behavioral Emphasis* (Human Kinetics, 2005), fourth ed. edn.
- [3] Rosenbaum, D. A. *Human Motor Control* (Elsevier, 2010).
- [4] Yarrow, K., Brown, P. & Krakauer, J. W. Inside the brain of an elite athlete: the neural processes that support high achievement in sports. *Nat Rev Neurosci* **10**, 585–596 (2009).
- [5] Dayan, E. & Cohen, L. G. Neuroplasticity subserving motor skill learning. *Neuron* **72**, 443–454 (2011).
- [6] Snoddy, G. S. Learning and stability: A psychophysical analysis of a case of motor learning with clinical applications. *J App Psych* **10**, 1–36 (1926).
- [7] Crossman, E. R. F. W. A theory of the acquisition of speed-skill. *Ergonomics* **2**, 153–166 (1959).
- [8] Newell, K. M. & Rosenbloom, P. S. Mechanisms of skill acquisition and the law of practice. In Anderson, J. R. (ed.) *Cognitive skills and their acquisition*, 1–55 (Lawrence Erlbaum Associates, 1981).
- [9] Heathcote, A., Brown, S. & Mewhort, D. J. The power law repealed: the case for an exponential law of practice. *Psychon Bull Rev* **7**, 185–207 (2000).
- [10] Bassett, D. S. & Bullmore, E. T. Small-world brain networks. *Neuroscientist* **12**, 512–523 (2006).
- [11] Bassett, D. S. & Bullmore, E. T. Human brain networks in health and disease. *Curr Opin Neurol* **22**, 340–347 (2009).
- [12] Bullmore, E. & Sporns, O. Complex brain networks: Graph theoretical analysis of structural and functional systems. *Nat Rev Neurosci* **10**, 186–198 (2009).
- [13] Smith, S. M. *et al.* Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage* **23**, 208–219 (2004).
- [14] Woolrich, M. W. *et al.* Bayesian analysis of neuroimaging data in FSL. *NeuroImage* **45**, S173–S186 (2009).
- [15] Bullmore, E. *et al.* Wavelets and statistical analysis of functional magnetic resonance images of the human brain. *Stat Methods Med Res* **12**, 375–399 (2003).

- [16] Bullmore, E. *et al.* Wavelets and functional magnetic resonance imaging of the human brain. *NeuroImage* **23**, S234–S249 (2004).
- [17] Brammer, M. J. Multidimensional wavelet analysis of functional magnetic resonance images. *Hum Brain Mapp* **6**, 378–382 (1998).
- [18] Achard, S., Salvador, R., Whitcher, B., Suckling, J. & Bullmore, E. A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs. *J Neurosci* **26**, 63–72 (2006).
- [19] Bassett, D. S., Meyer-Lindenberg, A., Achard, S., Duke, T. & Bullmore, E. Adaptive reconfiguration of fractal small-world human brain functional networks. *Proc Natl Acad Sci USA* **103**, 19518–19523 (2006).
- [20] Achard, S. & Bullmore, E. Efficiency and cost of economical brain functional networks. *PLoS Comput Biol* **3**, e17 (2007).
- [21] Achard, S., Bassett, D. S., Meyer-Lindenberg, A. & Bullmore, E. Fractal connectivity of long-memory networks. *Phys Rev E* **77**, 036104 (2008).
- [22] Bassett, D. S., Meyer-Lindenberg, A., Weinberger, D. R., Coppola, R. & Bullmore, E. Cognitive fitness of cost-efficient brain functional networks. *Proc Natl Acad Sci USA* **106**, 11747–11752 (2009).
- [23] Lynall, M. E. *et al.* Functional connectivity and brain networks in schizophrenia. *J Neurosci* **30**, 9477–9487 (2010).
- [24] Percival, D. B. & Walden, A. T. *Wavelet Methods for Time Series Analysis* (Cambridge University Press, 2000).
- [25] Sun, F. T., Miller, L. M. & D’Esposito, M. Measuring interregional functional connectivity using coherence and partial coherence analyses of fMRI data. *NeuroImage* **21**, 647–658 (2004).
- [26] Bassett, D. S. *et al.* Dynamic reconfiguration of human brain networks during learning. *Proc Natl Acad Sci USA* **108**, 7641–7646 (2011).
- [27] Mucha, P. J., Richardson, T., Macon, K., Porter, M. A. & Onnela, J.-P. Community structure in time-dependent, multiscale, and multiplex networks. *Science* **328**, 876–878 (2010).
- [28] Porter, M. A., Onnela, J.-P. & Mucha, P. J. Communities in networks. *Not Amer Math Soc* **56**, 1082–1097, 1164–1166 (2009).
- [29] Fortunato, S. Community detection in graphs. *Phys Rep* **486**, 75–174 (2010).
- [30] Newman, M. E. J. & Girvan, M. Finding and evaluating community structure in networks. *Phys Rev E* **69**, 026113 (2004).
- [31] Newman, M. E. J. Modularity and community structure in networks. *Proc Natl Acad Sci USA* **103**, 8577–8582 (2006).
- [32] Newman, M. E. J. Finding community structure in networks using the eigenvectors of matrices. *Phys Rev E* **74**, 036104 (2006).
- [33] Jutla, I. S., Jeub, L. G. S. & Mucha, P. J. A generalized Louvain method for community detection implemented in MATLAB (2011–2012). URL <http://netwiki.amath.unc.edu/GenLouvain>.

- [34] Blondel, V. D., Guillaume, J. L., Lambiotte, R. & Lefebvre, E. Fast unfolding of community hierarchies in large networks. *J Stat Mech* P10008 (2008).
- [35] Rombach, M. P., Porter, M. A., Fowler, J. H. & Mucha, P. J. Core-periphery structure in networks. *arXiv:1202.2684* (2012).
- [36] Bollobás, B. *Modern Graph Theory* (Springer Verlag, 1998).
- [37] Dodge, Y. *The Oxford Dictionary of Statistical Terms* (Oxford University Press, 2003).
- [38] Darlington, R. B. Is kurtosis really ‘peakedness’? *The American Statistician* **24**, 19–22 (1970).
- [39] Mesulam, M. M. From sensation to cognition. *Brain* **121**, 1013–1052 (1998).
- [40] Newman, M. E. J. Fast algorithm for detecting community structure in networks. *Phys Rev E* **69**, 066133 (2004).
- [41] Bassett, D. S. *et al.* Robust detection of dynamic community structure in networks. *arXiv:1206.4358* (2012).
- [42] Simpson, S. L., Moussa, M. N. & Laurienti, P. J. An exponential random graph modeling approach to creating group-based representative whole-brain connectivity networks. *NeuroImage* **60**, 1117–1126 (2012).
- [43] Traud, A. L., Kelsic, E. D., Mucha, P. J. & Porter, M. A. Comparing community structure to characteristics in online collegiate social networks. *SIAM Rev* **53**, 526–543 (2011).
- [44] Guimerà, R., Sales-Pardo, M. & Amaral, L. A. Classes of complex networks defined by role-to-role connectivity profiles. *Nat Phys* **3**, 63–69 (2007).
- [45] Peters, R. A. *Biochemical lesions and lethal synthesis* (Pergamon Press, Oxford, 1963).
- [46] Clauset, A., Shalizi, C. R. & Newman, M. E. J. Power-law distributions in empirical data. *SIAM Review* **51**, 661–703 (2009).
- [47] Stumpf, M. P. H. & Porter, M. A. Critical truths about power laws. *Science* **51**, 665–666 (2012).
- [48] Hagmann, P. *et al.* Mapping the structural core of human cerebral cortex. *PLoS Biol* **6**, e159 (2008).
- [49] Bassett, D. S., Brown, J. A., Deshpande, V., Carlson, J. M. & Grafton, S. T. Conserved and variable architecture of human white matter connectivity. *NeuroImage* **54**, 1262–1279 (2011).
- [50] Measuring signal power (2012).
- [51] Smith, S. W. *Digital Signal Processing: A Guide for Engineers and Scientists* (California Technical Pub, 1997).
- [52] Lazar, N. *The Statistical Analysis of Functional MRI Data* (Springer, 2010).
- [53] Grinband, J., Wager, T. D., Lindquist, M., Ferrera, V. P. & Hirsch, J. Detection of time-varying signals in event-related fMRI designs. *Neuroimage* **43**, 509–20 (2008).
- [54] Steffener, J., Tabert, M., Reuben, A. & Stern, Y. Investigating hemodynamic response variability at the group level using basis functions. *NeuroImage* **49**, 2113–2122 (2010).

- [55] Calhoun, V. D., Adali, T. & Pekar, J. J. A method for comparing group fMRI data using independent component analysis: application to visual, motor and visuomotor tasks. *Magn Reson Imaging* **22**, 1181–1191 (2004).
- [56] Reichardt, J. & Bornholdt, S. Statistical mechanics of community detection. *Phys Rev E* **74**, 016110 (2006).
- [57] Sheppard, J. P., Wang, J. P. & Wong, P. C. Large-scale cortical network properties predict future sound-to-word learning success. *J Cogn Neurosci* **24**, 1087–1103 (2012).
- [58] Whitlow, C. T., Casanova, R. & Maldjian, J. A. Effect of resting-state functional mr imaging duration on stability of graph theory metrics of brain network connectivity. *Radiology* **259**, 516–524 (2011).
- [59] Kramer, M. A. *et al.* Emergence of persistent networks in long-term intracranial eeg recordings. *J Neurosci* **31**, 15757–15767 (2011).
- [60] Chu, C. J. *et al.* Emergence of stable functional networks in long-term human electroencephalography. *J Neurosci* **32**, 2703–2713 (2012).
- [61] Jones, D. T. *et al.* Non-stationarity in the “resting brain’s” modular architecture. *PLoS One* **7**, e39731 (2012).
- [62] Siebenhühner, F., Weiss, S. A., Coppola, R., Weinberger, D. R. & Bassett, D. S. Intra- and inter-frequency brain network structure in health and schizophrenia. *In Preparation* (2012).
- [63] Barnes, A., Bullmore, E. T. & Suckling, J. Endogenous human brain dynamics recover slowly following cognitive effort. *PLoS One* **4**, e6626 (2009).
- [64] Tambini, A., Ketz, N. & Davachi, L. Enhanced brain correlations during rest are related to memory for recent experiences. *Neuron* **65**, 280–290 (2010).
- [65] Shanahan, M. & Wildie, M. Knotty-centrality: finding the connective core of a complex network. *PLoS One* **7**, e36579 (2012).
- [66] Uddin, L. Q., Supekar, K. S., Ryali, S. & Menon, V. Dynamic reconfiguration of structural and functional connectivity across core neurocognitive brain networks with development. *J Neurosci* **31**, 18578–18589 (2011).
- [67] Borgatti, S. P. & Everett, M. G. Models of core/periphery structures. *Social Networks* **21** (1999).
- [68] Holme, P. Core-periphery organization of complex networks. *Phys Rev E* **72**, 046111 (2005).
- [69] da Silva, M. R., Ma, H. & Zeng, A. P. Centrality, network capacity, and modularity as parameters to analyze the core-periphery structure in metabolic networks. *Proceedings of the IEEE* **96**, 1411–1420 (2008).
- [70] Garas, A., Schweitzer, F. & Havlin, S. A k-shell decomposition method for weighted networks. *arXiv:1205.3720* (2012).
- [71] Colizza, V., Flammini, A., Serrano, M. A. & Vespignani, A. Detecting rich-club ordering in complex networks. *Nature Physics* **2**, 110–116 (2006).
- [72] Opsahl, T., Colizza, V., Panzarasa, P. & Ramasco, J. J. Prominence and control: the weighted rich-club effect. *Phys Rev Lett* **101**, 168702 (2008).

- [73] van den Heuvel, M. P. & Sporns, O. Rich-club organization of the human connectome. *J Neurosci* **31**, 15775–15786 (2011).
- [74] van den Heuvel, M. P., Kahn, R. S., Goñi, J. & Sporns, O. High-cost, high-capacity backbone for global brain communication. *Proc Natl Acad Sci USA* **109**, 11372–11377 (2012).
- [75] Ekman, M., Derrfuss, J., Tittgemeyer, M. & Fiebach, C. J. Predicting errors from reconfiguration patterns in human brain networks. *Proc Natl Acad Sci U S A* **Epub Ahead of Print** (2012).
- [76] Sepulcre, J., Sabuncu, M. R., Yeo, T. B., Liu, H. & Johnson, K. A. Stepwise connectivity of the modal cortex reveals the multimodal organization of the human brain. *J Neurosci* **32**, 10649–10661 (2012).
- [77] Fornito, A., Harrison, B. J., Zalesky, A. & Simons, J. S. Competitive and cooperative dynamics of large-scale brain functional networks supporting recollection. *Proc Natl Acad Sci USA* **109**, 12788–12793 (2012).
- [78] Dosenbach, N. U. *et al.* A core system for the implementation of task sets. *Neuron* **50**, 799–812 (2006).
- [79] de Pasquale, F. *et al.* A cortical core for dynamic integration of functional networks in the resting human brain. *Neuron* **74**, 753–764 (2012).